


```

170 TyrGluLysAlaGlyValAlaProPheAlaGluGluAspAspProAlaGlu 186
531 ATACGGCTGCGCTTGTGTACACTCAAGCAGGCTTCAACCAAGACAGAGA 580
186 uTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCATGGCAGACAGCTGCTGCGGCGCATGAACGCTTTCCGAGAGACA 630
203 spLysTrpThrLysIleAlaGluSerValLysGlyValThrGluGluThr 219
631 ACAACAGGTGTCACCGCTCTACAGCTCGACGTCGAGAGAGAGGCAACTCT 680
220 ThrThrGlyValLeuArgLeuArgLeuGlnPheAlaAlaGlyAspLeuAl 236
681 CTTCACAGCCATCAACGTCACAGCAGCTGTTCACAAAGTCCAAATCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGCGCTGCGCCACCTCCCTATGATGATGATCAACGCTGCTCC 780
253 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATCGCGGCGCAAGACAGCTCTGTCATGCGGTACGGCATGT 830
270 AspAlaLeuIleGlyGlyLysLysValLeuIleCysGlyTyrGlyAspVa 286
831 CGGCAGAGGCTGCGCTCAATCCCTCGTGGCCAGAGGCTCGGCTATCA 880
286 lGlyLysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTGACCCAACTGCGCTCTCCAGGCTCCCATGGAAGGCTAC 930
303 alThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGGTCCGCGCCATCGAGCAAGTGGTCAAGATGTCATATCTTCCTTAC 980
320 AspAlaValThrValGluGluAlaIleGlyAspAlaAspIleValAlaTh 336
981 ATGACAGAGAACTGCGATATCATCTGTGATCATGATGATGCGCCAGATGA 1030
336 rAlaThrGlyAsnLysAspIleIleMetLeuGlnHisIleLysAlaMetL 353
1031 AGGATTAAGGCTATGTGCGTAAACATCGCCACTTGATACCAAAATTGAT 1080
353 yAspHisAlaIleLeuGlnLysnIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGGCCTCATGAATATACCCAGCATCAAGCATCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysP 385
1131 AGAATACGACATGTGGGAATTCACAGAT...GGCCAGGCTATCCTCTTC 1177
385 oGlnValAspLeuTrpThrPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGGTGAGGCGCGCTTCTTAACCTGGCTGGCTGACAGAGTCAACCAGTCT 1227
402 euSerGluGlyArgLeuLeuAsnLeuGlyAsnIleThrGlyHisProSer 418
1228 TTGCTTATGTCATGTCTACAAACAGACACTCGCTCAGCTGACAGCT 1277
419 PheAlaMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTACGAAAGAGAGGAAATCTCGAGAGAAAGGTTTACACTTCCGAGAC 1327
435 uTrpThrLysAsnAspGluTyrAspAsnGluValTyrArgLeuProLysH 452
1328 ATCTCGATGAAGAGCTGCGCTCCACCTCGGATCTCTCGATGCTCCAC 1377
452 lIleuAspGluLysValAlaIleArgIleHisValGluAlaLeuGlyLysH 468
1378 CTTACAAAGCTTACACAGAGCAGGCTGACTACATCAACGTTCCAGTTTA 1427

```

```

469 LeuThrLysLeuThrLysGluGlnAlaGluTyrLeuGlyValAspValAl 485
1428 GGTCCTTACAGTGTGATGCTTACCGTAT 1458
485 uGlyProTyrLysProAspHisTyrArgTyr 495

```

seq_name: /cgn2_6/ptodata/2/paa/US097_COMB.pep:US-09-708-427-22304

seq_documentation_block:

; Sequence 22304, Application US/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEP

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 22304

; LENGTH: 485

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1..485

; OTHER INFORMATION: Xaa is any amino acid

; NAME/KEY: misc_feature

; LOCATION: 1..485

; OTHER INFORMATION: Ceres Seq. ID 1840722

; US-09-708-427-22304

alignment_scores:

Quality: 1537.00 Length: 487

Percent Similarity: 80.903 Percent Identity: 62.012

alignment_block:

US-09-759-990-1 x US-09-708-427-22304 ..

Align seg 1/1 to: US-09-708-427-22304 from: 1 to: 485

```

19 ACTGGTGTCTCATTTGAGTACAGAAATGCCGACATCAACCTCCATGTTCT 68
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspH 24
69 CGGCGGTAAAGAACTTACCTTGTCTGAGAAGAAATGCCAGGTTTATGG 118
24 eGlyArgLeuGlnLeuGlnLeuAlaGluValGluMetProGlyLeuMetA 41
119 TTCTTCGTGAGAGCTTATTCGCTTCTTAAGCCATTGAAGGTCAGAGAT 168
41 lacyArgThrLurPheGlyProSerGlnProPheLysGlyAlaIle 57
169 TCTGTTCCCTCCACATGACAGTCCAGACAGCCGCTCTCATGACAGACT 218
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACACCTTGTGTGCTGATGTCAGATGGCTTCTCTGACAACTTCTCTTA 268
74 uThrAlaLeuGlyAlaGluValArgTyrPcySerCysAsnIlePheSer 91
269 CACAAGATACAGCCGCTGCTGATGTCAGTGTGTCGCGCCCAACGACAGCA 318
91 hrgLysAspHisAlaAlaAlaIleAlaIleArgAspSerAlaAla..... 105
319 GAGAACCAAGCCGCTATCCAGTCTTCCGCTGAGAGCGCAACACATCCC 368
106 .....ValPheAlaTrpLysGlyGlnThrLeuGlu 115
369 AGAATACGGGAGAACACATACCCTGCTCTACATGCGCCAGATGTCAG 418

```

```

115 nglutrtprpcythrghluaralaleuasprpbglyproglylylg 132
419 GCCACACAGAGTTGTCGATGATGGTGGTACATCCATCC 468
132 lpproaspheulevalaspasglylyaspalatnleuilehis 148
469 AAGGCTTCGAA.....TTCGAAACAGCCGCTGCTGTCAGA 506
149 gluglyvalylaglugluilepneghlulysthrghlvalproas 165
507 GCCACAGAGAGTGCACACCTGCATACCCGCTTCTTCTACACICA 556
165 pprothserthrpspnproglupneghlilevalleuilerleiel 182
557 AGCAGCTCTCAACCAAGACAGACACAGCCAGCAGCTGCTCCGCG 606
182 ysgluglyleuglnvalasproullystlyhislysmellysluarq 198
607 ATGACGCGTGTTCGAGAGACACACAGCTGTCACCGCTTACCA 656
199 leuvaliglyvalsergluutntrthrghlyvalylasargleutygl 215
657 GCTCGAAGAGAGGCGCAACTCTCTTCCAGCATCAACGTCACAGAG 706
215 nmetglnghlansglythrleuupheproalaileasnvalasnaps 232
707 CHTTACAAGTCCAGTTCATACATCTACGCGCTCCGCTCCCTT 756
232 eivalthrlysserlyspneaspnleutyrglycysarghisserleu 248
757 ATCGATGTATCAACCGTCTCCGATGTATGATGGCGGACAGACAGC 806
249 proaspglyleumetaralathraspvalmetilealaglylyvalal 265
807 TCTCGTCATGGTTCAGCGCATGTCGCAAGGCGCGCTCATCCCTCC 856
265 aavalliecygsltyrglyaspvaliglylysglyalaalaleumel 282
857 GTGGCAAGGCGCTCGCTTATCATCAAGACCACTGACCCAACTTCGCT 906
282 ysthrilaglyalargvalillevalthrghlileasproillecysala 298
907 CTCGAGGCTGCGATGGAAGGCTACCAAGTCCGCGCATCGAGAACTCGT 956
299 leughlalaileumetgluglyleughlvalleuthrleuighlaspvalya 315
957 CAAGGATTCGATATCTTCCTTACATGACAGAAACTCGCATATCATCT 1006
315 lserghlualaspillepnevalthrthrghlyasnlyasbillelem 332
1007 CTGTTGCATGATGCCAGATGAAGATTAAGCTATTGTCGGTAAACATC 1056
332 etvalaspnismetarglysmetlyasnasnallevalcysasnille 348
1057 GGCACACTGCATACGAATATGATACAGATGGCTCATGAATACCAGG 1106
349 glyhisphespnasngluileaspmetleuglyleughluthrtyrprogl 365
1107 CATCAACCAATCCCAATCAAGCCAGAAATACATGATGGGAATTCACAG 1156
365 yvallysargllethrlelyspnlethrthrparagtrpvalpneprog 382
1157 ATGGCCAGCT...ATCCCTCTTCTGCTGAGGCGCGCTTCTTAACTT 1203
382 luthrlysalaglyllelevalleualagluglyargleumetasnleu 398
1204 GCGTGGCTCAGATCAGCCATCTTCTGTATGATCAATGATTCACAAA 1253
399 glycysalatrghlyhisproserpnevalmetsercysserpnehras 415
1254 CCAGACACTCGCTCAGCTCGACCTTCAGCAAGAAAGAGA.....GGAATC 1297
415 nglvalillealaglnleuglnleutrpasnghlulysalaserglyyst 432

```

```

1298 TCGAAGAGAGTTTACACACTTCCGAAAGCATTCGATGAAGAGTCTCT 1347
432 yrglulyslysaltyrvalleuproullyshisleuaspglulysvalala 448
1348 CGCCTCCACCTCGGATCTCTCGATGTCCACCTTCAAGCTTACACAGA 1397
449 leuilehisleuileuglylsleuglyalargleuthrlysserlyas 465
1398 GCAGGCTGACTACATCAACCTTCAGTTCAGAGGCTCTTCAAGTGTGATG 1447
465 pmlnseraptyrvalserleleproleleuglylprotyrlyspnprofn 482
1448 CTTACCGTTAT 1458
482 lstryargtyr 485

```

```

seq_name: /cgn2_6/ptodata/2/pae/US60_COMB.pap:US-60-361-294-16
seq_documentation_block:
; Sequence 16, Application US/60361294
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: McElver, John Alan
; APPLICANT: Zhou, Qing
; APPLICANT: Aux, George W.
; APPLICANT: Tossberg, John
; APPLICANT: Ashby, Carl
; APPLICANT: Dunn, Jill
; APPLICANT: Cates, Eddie
; APPLICANT: Law, Marcus Dixon
; APPLICANT: Budzieszewski, Greg
; APPLICANT: Nyfeler, Beat
; APPLICANT: Wu, Jia-Qian
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential For Plasmid Replication
; FILE REFERENCE: 60169P1
; CURRENT APPLICATION NUMBER: US/60/361,294
; CURRENT FILING DATE: 2002-02-28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 485
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
US-60-361-294-16

```

```

alignment_scores:
Quality: 1537.00 Length: 487
Ratio: 3.901 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 62.012

```

```

alignment_block:
US-09-759-990-1 x US-60-361-294-16 ..

```

```

Align seg 1/1 to: US-60-361-294-16 from: 1 to: 485

```

```

19 ACTGCTGCTCATTCGATGACGAAATGGCGACATCAACCTCATGTTCT 68
8 Thrserserghlyargglutryllysalylaspmetserlnalaspn 24
69 CGGCGCTGAGAACTTACCTTGCCTGAGAAAGAAATCCAGCTTATG 118
24 egllyargleughluluglnleualagluvalglumetprouglyleumeta 41
119 TCTTCGTGAGCGTATTCCTGCTTCAAGCATGAGGCTGTCAGATC 168
41 lacysarghthrcluhneghlyproserclnproheuglyalargyle 57
169 TCTGCTCCCTTCACATGACAGTCCAGACAGCGCTCTCATCGAGACT 218

```

```

58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTGTGTGCTGATGATGAGTGGCTCTCCGCAACATCTTCTCTA 268
|||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
269 CACAGATACAGCGCGTGTGATGCTGTGTGCGGCCAACAGGACCA 318
|||||
91 hGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAAGCCAGCGGTATCCAGTCTTCGCTGGAAAGGCGCAACACCTCC 368
|||||
106 .....ValPheAlaIleTrpIleGluValLeuIleGlu 115
369 AGAATACCTGGAGAACACATACCGCGCTCTCAGATGGCCAGATGCTAAG 418
|||||
115 nGluTrpTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 132
419 GCGCCAGACAGGTGTGATGATGATGATGATGATGATGATGATGATGAT 468
|||||
132 LyrProAspLeuIleValAspAspGlyGlyAspAlaThrLeuIleHis 148
469 AAGGCTTCGAA.....TTCGAAACAGCGGTGCTGTTCGAA 506
|||||
149 GluGlyValIleGluGluIlePheGluIleGluValIleGluValProAs 165
507 GCCAACACAGCTGACACCTGGAATACCGCTCGCTTGTGCTACACTCA 556
|||||
165 pProHisSerTrpAspAsnProGluPheGlnIleValLeuSerIleIle 182
557 AGCAGGTCTTCACCAACAGAACACACACACACACACACACACACACAC 606
|||||
182 YsgLuglyLeuGlnValAspProIleGlySerIleGlyValIleGluArg 198
607 ATGAACGCTGTTCGAAAGAGACACACACACACACACACACACACACAC 656
|||||
199 LeuValGlyValSerGluIleThrThrThrGlyValIleArgLeuTrpGlu 215
657 GCTGAGAGAGAGGCGCAACCTCTCTCCAGCAACACACACACACACACAC 706
|||||
215 nMeGlnIleAsnGlyThrLeuLeuPheProAlaIleAsnValAspAsp 232
707 CTGTTCACCAAGTCCAGTTCATACATCTACGCGTGGCGGCGCTGCTT 756
|||||
232 eValIleTrpSerTrpPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
757 ATGATGATGATCAACCGCTGCTTCGATGATGATGATGATGATGATGAT 806
|||||
249 ProAspIleLeuMetArgAlaThrAspValMetIleAlaGlyValValAl 265
807 TCTGCTATGAGGTACGCGGATGTCGCAAGGCGCGCTCAATCCCTCC 856
|||||
265 aValIleCysGlyTrpGlyAspValGlyGlyGlyAlaAlaAlaMet 282
857 GTGGCCAGGCGCTGCGCTTATCATACACAGACTGACCCCAATCTGCGT 906
|||||
282 YsThrAlaGlyAlaArgValIleValIleThrGluIleAspProIleCysAla 298
907 CTCGAGGCTGCGATGAGAGCTTACAGGCTCGCCCATCGAGGAAGTCTG 956
|||||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValAla 315
957 CAAGATGTCGATATCTGTTACATGACAGAAATGCGCATATCATCT 1006
|||||
315 LserGluAlaAspIlePheValIleThrThrGlyAsnIleGlyAspIleIle 332
332 eValIleAspHisMetArgIleGlyMetIleGlyAsnAlaIleValCysAsnIle 348
1057 GGCCACTTCGATACGAAATGATACAGATGCGCTCATGAATACCCAGG 1106
|||||
349 GlyHisPheAspAsnGluIleAspMetIleGlyLeuGluTrpTrpGlu 365

```

```

1107 CATCAAGCACATCCCAATCAAGCCAGAAATACACATGTGGATTTCCAG 1156
|||||
365 YValIleArgIleThrIleIleGlySerGlnThrAspArgTrpValPhePro 382
1157 ATGGCCAGCGT...ATCCTCCTTCTGCTGAGGCGCGCTTCTTAACCTT 1203
|||||
382 LurThrIleAlaGlyIleIleValLeuAlaGluGluArgLeuMetAsnLeu 398
1204 GCGTGGCTACAGGTACACCCATCTTGTGATATGTCATGTCACAAA 1253
|||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACCTGCGTCAGCTCGACCTTACGAAAGAGA.....GGAATTC 1297
|||||
415 nGluValIleAlaGlnIleGluIleuTrpAsnGluIleValAspGlyVal 432
1298 TCAGAGAAAGGTTTACACACTTCGAAAGCATCTGATGATGAGAAATGCT 1347
|||||
432 YrGluIleGlyValIleValLeuProIleGlyHisIleuAspGluIleValAla 448
1348 CGGCTCCACCTCGATCTCTGATGTCACCTTACAAAGCTTACACAGAA 1397
|||||
449 LeuIleuHisIleuGlyIleuGlyAlaArgLeuThrIleuSerIleuAs 465
1398 GCAGGCTGATCATCAACGTTCCAGTTGAGGTCCTTACAAAGTGTGATG 1447
|||||
465 pGlnSerAspTrpValSerIleProIleGluGlyProIleIleuSerProIle 482
1448 CTTTACCGTTAT 1458
|||||
482 LsTyArgTrpTyr 485

```

```

seq_name: /cgn2_6/ptodata/2/pae/US097_COMB.pep:US-09-708-427-77520
seq_documentation_block:
? Sequence 77520, Application US/09708427
? GENERAL INFORMATION:
? APPLICANT: N. ALEXANDROV et al.
? TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEP
? FILE REFERENCE: 2750-1243P
? CURRENT APPLICATION NUMBER: US/09/708,427
? CURRENT FILING DATE: 2000-11-09
? NUMBER OF SEQ ID NOS: 85364
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 77520
? LENGTH: 485
? TYPE: PRT
? ORGANISM: Zea mays subsp. mays
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1..485
? OTHER INFORMATION: Xaa is any amino acid
? NAME/KEY: misc_feature
? LOCATION: 1..485
? OTHER INFORMATION: Ceres Seq. ID 1960872
US-09-708-427-77520

```

```

alignment_scores:
  Quality: 1523.00      Length: 487
  Ratio: 3.875          Gaps: 4
  Percent Similarity: 80.698  Percent Identity: 61.807

```

```

alignment_block:
US-09-755-990-1 x US-09-708-427-77520 ..

```

```

Align seg 1/1 to: US-09-708-427-77520 from: 1 to: 485

```

```

19 ACTGATGCTCATTCGAGTACAGAAATGCGCAACATCACTCATTTCT 68
|||||
8 ThrSerGlyArgGluTrpIleValIleAspLeuSerGlnAlaAsp 24

```



```

69 CGGCGGTAAAGCACTTACCCCTGCTGAGAAAGAAATGCCAGCTTATGG 118
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 eglYArgLeuGluIleGluLeuValGluMetProGlyLeuMetA 41
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TTTCTGCGAGGGTATGCCGTTTCAAGCAATGAAGGGTGTCAAGTC 168
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 IAcysArgAlaIuPheGlyProSerLysProPheAlaGlyAlaArgIle 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TCTGTTCCCTCCACATGACAGTCCAGACAGCCGCTCATGAGACACT 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 CACAGCTTTGGTGTGATGATGACAGATGGCTCTCCGCAACATCTTCTA 268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerT 91
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CACAGATACACAGCCGCTGCTGATGCTGTGCGGCCACAGCAGCACCA 318
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 hrgInAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 GAGAAGCCAGCCGGTATCCAGTCTTCGCTGAGAGGCGCAACACTCC 368
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 .....ValPheAlaTrpLysGlyIuThrLeuGlu 115
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 AGAATACTGGAGAACACATACCAGCGCTTCACATGGCCAGATGGTCAAG 418
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 uGluTrpTrpTrpCysThrGluArgCysLeuAspTrpGlyGluAlaGlyG 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 GCCCAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 468
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 LProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
469 AAGGGCTTC.....GAATTCGAAACAGCCGGTCTGTTCAGAA 506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GluGlyValIysAlaGluGluGluGluGluGluGluGluGluGluGlu 165
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 GCCACAGACAGTGCACACCTGCATACCCGCTGCTTCTGCTACATCA 556
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 pProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleIleA 182
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 AGCAGGTTCTCAACCAAGAACCAACCAACCAACCAACCAACCAACCA 606
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 rGaSPglYleuYsAlaSPProLysIlyTrGlyMetLysGluArg 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 ATGAACGGTGTTCGAGAGACAAACAAACAGTGTCCACGCTCTACCA 656
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 LeuValGlyValSerGluGluGluGluGluGluGluGluGluGluGlu 215
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 GCTCGAGAGAGGCGCAACTCCTCTCCAGCCATCAACGTCACGACG 706
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 mEtGlnGluThrGlyAlaLeuPheProAlaIleAsnValAsnAspS 232
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 CTGTTCACAAGTCCAAAGTTCGATACATCTACGCTGCCGCACTCCCT 756
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 eValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 ATCGATGTATCAACCGCTCTCCAGTGCATGATCGCGCGCAGACAGC 806
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 TCTGCTCATGGGTACGGCGATGTCGCAAGGCTGCGCTCAATCCCTCC 856
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 aValValCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaAlaLeu 282
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 GTGGCCAGAGGCGCTGCGCTTATCATCAAGACGACGACGACGACGCT 906
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 ySglnAlaGlyAlaArgValIleValThrIuIleAspProIleCysAla 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 CTCGAGGCTGCATGAGAGCTACAGGTCGCGCATCGAAGAGTGTGT 956
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuGlnAlaLeuMetGluGluGluGluGluGluGluGluGluGluGlu 315

```

```

957 CAAGATGTCGATATCTTCTGATACATGCACAGGAACCTGCCATATCATCT 1006
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 lSerIuAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1007 CTGTGACATGATGGCCAGATGAAGGATGAAGGATATGTGCTGCTAACC 1056
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 eValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1057 GCCCACTTCGATACAGAAATTTGATACAGATGGCCCTCATGAATACCAAG 1106
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
349 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpProGlu 365
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1107 CATCAAGCAATCCCAATCAAGCCAGAAATACAGATGTGGGAATTTCCAG 1156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 yAllyAspArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1157 ATGGCCAC...GCTATCCCTCTTCTGATGAGGCGCGCTCTTAACCTT 1203
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 luThrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1204 GGCTGCGCTTACAGGTACACCATCTTCTGATATGATATGATATTCACAAA 1253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1254 CCAGACACTGCTGACGCTGACCTCTACGAAAGA.....GGAATC 1297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 nGlnValIleAlaGlnLeuGluLeuTrpLysGluLysSerSerGlyLysT 432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1298 TCGAGAGAGAGGTTTACACACTCCGAAAGATTCGATGAGAGAACTGGT 1347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
432 yGluLysLysValTrpValLeuProLysHisLeuAspGluLysValAla 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1348 CGCCTCCACCTGGATCTTCGATGTCCACCTTCAAAAGCTTACACAGAA 1397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
449 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrLysSe 465
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1398 GCAGGCTGACTATCATCAAGCTTCAGTGAAGGTCCTTCAAGTGTATG 1447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 rGlnAlaAspTrpLysSerValProIleGluGlyProTrpLysProAlaH 482
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1448 CTTACCGTTAT 1458
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
482 lSTyTrpGlyTr 485
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: /cgn2_6/ptodata/2/paa/us097_COMB.pep:us-09-708-427-77519

seq_documentation_block:

```

; Sequence 77519, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPE
; FILE REFERENCE: 2750-1243p
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77519
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..510
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..510
; OTHER INFORMATION: Ceres Seq. ID 1960871
; US-09-708-427-77519

```

alignment_scores: Quality: 1523.00

Length: 487

Ratio: 3.875 Gaps: 4
Percent Similarity: 80.698 Percent Identity: 61.807
Alignment_block:
US-09-759-990-1 x US-09-708-427-77519 ..
Align seg 1/1 to: US-09-708-427-77519 from: 1 to: 510

```

19 ACTGTCCTCCATCGAGTACAGATTGCGCAATCAACCTTCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 Thrserserlglyargglutrylsvalylsaspleserlnalaspsh 49
69 CGGCCCGTAAGAACTTACCTTCTGAGAAAGAAATGCCAGCTTTATG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
49 eglvalgluugluilegluleuallagluvalglumetproglyleu 66
119 TTCCTGTCGAGCGTTATCCGCTCTTACGCAATTAAGAGCTGCAATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
66 lacsarlgallagluipheglyproserlysprophelaglyalarg 82
169 TCTGTCCCTCCACATGACATGACAGCCGAGCCGCTCATGAGACACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 SerglyserleuHismetThrileglthrAlaValleuileglut 99
219 CACAGCTTGTGCTGATGTCAGATGGGCTTCTGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
99 utrlalaleuglyalagluvalargtrpCysserCysasnilephes 116
269 CACAGATACAGCGCTGCTGTATGCTTGGCGCCCAAGGCAACAGCA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
116 hrlnaspshlsalaleualaleualargspserlala..... 130
319 GAGAAAGCCAGCCGGTATCCAGTCTTCGCTGAGAGGCGGAACACTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
131 .....Valpheelatrlpysglyluthrleucl 140
369 AGAATACTGGGAGAACACATACCGCGCTCTCATAGGCCAGTGGTCA 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
140 uglutrytrpCysThrIuargCysleuasprrpIgluallaglyg 157
419 GCCCAGCAGAGTGTGATGATGATGGTGTGATGATGATGATGATGAT 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
157 lypkaspplleuilevalaspaspplglyaspalathrleuileu 173
469 AAGGGCTTC.....GAATTCGAACAGCGGCTGCTTCAGCA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 GluIglYallysalagluIglutrylglutrylthrglylsilepro 190
507 GCCAAGCAGAGCTGACAACTCGAATACCGGCTTCTGCTACACTCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
190 pProgluIserThrAspaspnaIagluIphelYsilevalleuThr 207
557 AGCAGGTCTTCAACCAAGCAACCAAGCAACCAAGTGTCCAGCCGCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 rgsaspclyleuIysalaspProIysIysIyrarglysmetIysclua 223
607 ATGACGGTGTTCGAGAGACAAACACAGAGTGTCCAGCCGCTTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
224 leuvalIglYalsergluIglutThrThrIglYallYsargleuY 240
657 GCTCGAGAGAGAGGCAACTCCTTCCAGCCATCAACGTCACAGCAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
240 lmetIgluIglThrIglYalaleuIeupheProIalaleasnvala 257
707 CTGTTCACAAAGTCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 ervalthrIysserIysPheaspasnleuIyrIglYcysarIghIse 273
757 ATCGATGTATCAACCGCTGCTTCGATGTATGATGCGGCGGCAAGAC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 ProaspclyleuIeumetargalathrAspvalmetilealagly 290
807 TCTGTCATCGGTTACGGCGATGTCGAGCAAGCGCTGCTCATCCCT 856

```

```

290 aValvalCysglYtrYglYaspvalglYysglYcysalaleu 307
857 GTGGCAGAGCGCTGCGCTTATCATCACAGAACCTGACCCATCTGCG 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
307 ysglnalaglyAlarYvalIlevalThrIgluleaspProIleCysa 323
907 CTCAGGCTGCCATGAAAGGCTACAGAGTCCGCGCATGAGAGTGGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
324 leuIglalaleuIeumetgluIglYleuIglValleuProIeuglu 340
957 CAAGATGTGCATATCTTCGTACATGCACAGGAAGCTGCAATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1007 CTGTGACATGATGCGCCAGATGAAGATGAAGCTTATGCTGCTTAC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
357 etValaspshIsmetarglysmetIysaspasnaleuIeualCysa 373
1057 GGCACCTTCGATTAACGAATTGATACAGATGGCTTCATGAATACCA 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
374 glyIaspshaspasnIleaspmetleuglyleuIglutThrIyrp 390
1107 CATCAGCAGCATCCCAATCAACCCAGAAATACGATGTGGATTCGCA 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
390 YallYsarIglIethrIleIyspProIlnhrspargtrpvalp 407
1157 ATGGCCAC...GCTATCCCTCTTGTGTCGAGGCGCCCTTCAACTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
407 luthrAsnThrIglYleIleuIeualaleuIagluIyargleuIe 423
1204 GCGTCGCTACAGAGTACCCCATCTTTCGTTATGTCATGTCACAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
424 glyIysalathrIglYhisProserPheValmetSerCysserPhe 440
1254 CCAGACACTCCGCTCAGCTCGACCTGTACGAAAAAGAA..... 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
440 nglvalIlealaglnleuIgluleuItrpIysglYysSerSergly 457
1298 TCGAAGAGAGAGTTCACACTTCCGAACATCTCGATGAAAGAGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
457 YrgluIyIysvalIyvalIeupProIyIshIeunspglYysvala 473
1348 CGCCTCCACCTCGATCTCGATGTCCACCTTACAAAGCTTACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
474 AlaleuHIsleuIglYysleuIglYalalYsleuThrIysleu 490
1398 GCAGGTCGATCAATCAACAGCTTCAGTGAAGGCTCCTTCAACTG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
490 rglhalaspIyrlIeserValProIleIgluIglYrIglYsPro 507
1448 CTACCGCTTAT 1458
507 lstrYargTyr 510

```

seq_name: /cgn2_6/prodata/2/pae/US60_COMB.pep:US-60-324-109-25353

seq_documentation_block:
; Sequence 25353, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 25353
; LENGTH: 499

TYPE: PRT
ORGANISM: Zea mays
FEATURE:
US-60-324-109-25353

alignment_scores:
Quality: 1520.00 Length: 487
Ratio: 3.868 Gaps: 4
Percent similarity: 80.698 Percent identity: 61.602

alignment_block:
US-09-759-990-1 x US-60-324-109-25353 ..

Align seg 1/1 to: US-60-324-109-25353 from: 1 to: 499

```

19 ACTGGTCCATTCGATGACAGATTCGCGACATCCATCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
22 ThrSerSerGlyArgGluTrpLysValLysAspLeuSerGlnAlaAsp 38
CGCGCGTAAGAACTTACCTTGTGAGAAAGAAATGCCAGGCTTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 eGlyTrpLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMet 55
119 TTCTTCGTAGCGCTTATCCGCTTCTTAAGCCATTGAAGGCTCAGATC 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 IacYArGlaGluPheGlyProSerLysProPheAlaGlyAlaArgIle 71
169 TGTGTTCCCTCCACATGACATGACATGACAGACGCTCTCATGAGACACT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 88
219 CACAGCTCTTGTGTGATGTGATGATGGCTCTCTGACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 105
269 CACAAGATACAGCCGCTGCTGCTATCGTGTGCGCCCAACAGCACACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 hrGlnAspHisAlaAlaAlaAlaAlaAlaAlaArgAspSerAlaAla 119
319 GAGAAGCCAGCCGCTATCCAGCTCTGCGCTGGAAGGCGAAGAACTGCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 .....ValPheAlaTrpLysGlyLeuThrLeuGlu 129
369 AGAATTAAGTGGGAGACACATACCGCGCTCTACATGGCCAGATGGTCA 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 uGluTrpTrpTrpCysThrGluTrpGlyLeuAspTrpGlyAlaGlyG 146
419 GCCCAGACAGAGTGTGATGATGATGGTGTGATGCTACACTCTCTCAT 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 LyrProAspLeuIleValAspAspGlyAlaAlaThrLeuLeuIleHis 162
469 AAGGGCTTCT.....GAATGGAACAGCGCGTGTGTTCCACA 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 GluGlyValAlaLysAlaGluGluAspTrpGlyLysThrGlyLysIle 179
507 GCCAACAGAGAGTGTGATGATGATGATGATGATGATGATGATGATG 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 ProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleIle 196
557 AGCAGGTCTTCAACCAAGACAGACACTGCGACACAGTCTGCTGCCGC 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 rGAspGlyLeuLysAlaAspProLysLysTrpArgLysMetLysGlu 212
607 ATGAACGCTGTTCGAGAGACAAACAGAGTGTCCACCGGCTCTACCA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 LeuValAlaGlyValSerGluGluTrpThrThrGlyValLysArgLeu 229
657 GCTCGAAGAGAGGCAAACTCTCTCCAGCCATCAACGTCACAGCAG 706
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 nMetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAsnAs 246
707 CTGTTCACAAGTCCAACTGTGATACATCTACGGCTGCCGCCATCCCT 756

```

```

246 eValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 ATCGATGATTAACCGCTCTCCGATGCTCATGATCGCGGCAAGACAG 806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 279
807 TCTGTCATAGGCTTACGGCAGATGTCGGCAGAGGCTGCGCTCAAC 856
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 aValValCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaLeu 296
857 GTGGCCAAAGCGCTCGGCTTATCATCACAGAACTGACCCATCTGCG 906
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 yGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 312
907 CTCACGCTCCATGGAAGGCTACAGGCTCCGCGCCAGAGGAAGTCT 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuProLeuGluAspVal 329
957 CAAGATGTCGATATCTTCTTACATGCACAGAAATCGCATATCATCT 1006
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
329 lSerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleLeu 346
1007 CTGTTCACATGATGGCCAGATGACAGATTAAGCTATTGCTGTAATC 1056
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 eValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 362
1057 GGCCACTTCGATTAAGAAATGATACAGATGCGCTCATGAATACCCAG 1106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
363 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTrpPro 379
1107 CATCAAGCACATCCCAATCAAGCCAGAAATACGATGTGGAAATCCAG 1156
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 yAllyLysArgIleThrIleLysProGlnThrAspArgTrpValPhe 396
1157 ATGGCCAC...CGTATCTCTCTTCTGCTGAGAGCGCGCTTCAACT 1203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
396 lThrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 412
1204 GGCTCGCTACAGGTCAACCCATCTTCTGTTATGTCATGATTCACAA 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 429
1254 CCAGACATCGCTCAGCTCGACTCTACAGAAAGAGA.....GGAATC 1297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
429 nGluValIleAlaGluLeuGluTrpLysGlyLysSerGlyLys 446
1298 TCGAAGAAGAGTTTACACACTTCCGAAGCATCTGATGTAAGAAGTC 1347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 yGluLysLysValTrpValLeuProLysHisLeuAspGlyLysValAla 462
1348 CGCTCCAGCTCGATCTCTGATGTCACACTTACAAAGCTTACACAGA 1397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThr 479
1398 GCAGGCTACTATCAACGTTCCAGTTGAGGCTCTTACAAAGTGTGAT 1447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
479 rGlnAlaAspTrpLysLeuValProIleGluGlyProGlyLysProAla 496
1448 CTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 lAspTrpArgTrp 499

```

seq_name: /cgn2_6/plodata/2/paa/US60_COMB.pep:US-60-312-544-7754

seq_documentation_block:

; Sequence 7754, Application US/60312544

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Edgerton, Michael D

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong


```

seq_documentation_block:
; Sequence 26717, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324,109
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 26717
; LENGTH: 504
; TYPE: PRN
; ORGANISM: Zea mays
; FEATURE:
US-60-324-109-26717

```

```

alignment_scores:
  Quality: 1520.00      Length: 487
  Ratio: 3.868          Gaps: 4
  Percent Similarity: 80.698  Percent Identity: 61.602

```

alignment_block:

US-09-759-990-1 x US-60-324-109-26717 ..

Align seg 1/1 to: US-60-324-109-26717 from: 1 to: 504

```

19 ACTGGTGCCTCATTCAGATGACAGATGCGCAGATCCATCCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 ThrSerSerGlyArgGluTyrLysValLysAspLeuSerGlnAlaAsp 43
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGCGCTAGAGACTTACCTGCTGAGAGAAAGATGCGAGGTCTATG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMe 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCTCTGAGCGTATTCGCGCTTCTAAGCATTCAGAGGTGTCAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 LacysArgAlaGluPheGlyProSerLysProPheAlaGlyAlaArg 76
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGTCCTCCATGACATGACATGACATGACATGACATGACATGAC 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTGTCGTGATGTCAGATGGGCTTCGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 uThrAlaLeuGlyAlaGluValArgTyrPcySerCysAsnIlePheSer 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACAGCGCGTCTGCTATCGTTCGCGCCACAGGCAACCA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 hGlnAspHisAlaAlaAlaIleAlaAlaArgAspSerAlaAla..... 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAGAGCAGCGGATATCCAGTCTTCGCTGAGAGGCGAAACACTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 .....ValPheAlaTyrPcyGlyGluThrLeuGlu 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATAGTGGAGAGACATACCGCGCTCTCACATGCGCAGATGTCAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 uGluTyrTyrTrpPcySerThrGluArgCysLeuAspTyrGlyAlaGly 151
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GCCAGAGAGAGGTTCGATGATGGTGGTGGTGGTGGTGGTGGTGGT 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 LyrProAspLeuIleValAspAspGlyAlaAspAlaThrLeuLeuHis 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AAGGGCTTC.....GAATTCGAACACCGCGGTGCTGCTTCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
168 GluGluValLysAlaGluGluAspTyrGluLysThrGlyLysIleProAs 184

```

```

507 GCCAAGAGAGCTGACAACTCGAATACCGCTGCTTCTGCTACATCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 ProGluSerThrAspAsnAlaGluPheLysIleValLeuThrIleLeu 201
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 AGCAGGTCTTCACCAAGACAGAACACATCGGCGACAGTGTGCGCCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 rGAspGlyLeuLysAlaAspProLysLysTyrArgLysMetLysGluArg 217
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607 ATGAAGGCTGTTCCGAGAGACAGAACAGAGTTCACCGCTCTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 LeuValGlyValSerGluIleThrThrThrGlyValLysArgLeuTyrG 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657 GCTCGAAGAGAGGCGCAACTCTCTTCCAGCCATCAACCTCAACGAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 nMetGlnGluThrGlyAlaLeuPheProAlaIleAsnValAsnAspS 251
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707 CTGTACAAAGTCCAGTTCATCAATCTACGCTGCGCCGCTCAATCC 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 eValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 267
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757 ATCGATGTATCAACCGTCTCGATGTCATGATCGGCGGCAACAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 ProAspGlyLeuMetAlaArgAlaThrAspAlaMetIleAlaGlyValAl 284
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807 TCTCGTCATGAGGTTCAGCGCGCATGCGCAAGGCGTGCCTCAATCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 aValValLysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaLeu 301
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
857 GTGGCCAGAGCGCTCGCTATATCAACAGAACTGACCCATCTGCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 yGlnAlaGlyAlaArgValIleValIleThrGluIleAspProIleCysAla 317
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907 CTCACAGCTGCGCATGAGAGGTACAGGTCCGCGCATCGAGAACTGCT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
318 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuProLeuGluAspValVa 334
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957 CAAGATGTGATATTCCTGTTACATGACAGAACTGACCATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334 lSerGlnAlaAspIlePheValIleThrThrGlyLysnLysAspIleIle 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1007 CTGTGACATGATGCGCCAGATGAGAAAGATTAAGCTATTGCTGATA 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 eValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 367
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057 GGCACCTGATTAACGAATGATGACATGCGCTCATGAATACCCAG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
368 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProG 384
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107 CATCAAGCACATCCCAATCAAGCCGAAATACGACATGGGAATTCAC 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157 ATGGCCAC...GCTATCTCTCTCTGCTGAGGCGCGCTTTCATCCT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 lThrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 417
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204 GCGTCCGCTACAGTCAACCATCTTCTGATGTCATGATGATTCACAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheMet 434
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254 CCAGACACTCGCTGCTGACATCTACGAAAGAGAGAGAGAGAGAGAG 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 nGlnValIleAlaGlnLeuGluLeuThrLysGlyLysSerSerGlyLys 451
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298 TCGAGAGAGAGGTTCACACTTCGAGACATCTGATGAGAGAGTGGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 yGluLysLysValLysValLeuProLysHisLeuAspGlyLysValAla 467
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348 CGCTCCAGCTCGAGATCTCGATGTCGACCTTACAAAGTTCACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
468 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrLys 484
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398 GCAGGCTGACTACATCAACGTTCCAGTTCAGGCTCTTACAACTGTATG 1447

```

```

|||||
484 rGINAlAspIYrIleSerValProIleGIuGIyProIYrIleYsProAlaH 501
1448 CTACCGTTAT 1458
|||||
501 ISYrArGTyr 504

```

```

seq_name: /cgn2_6/prodata/2/paa/US60_COMB.pep:US-60-324-109-32880

```

```

seq_documentation_block:

```

```

; Sequence 32880, Application US/60324109
; GENERAL INFORMATION:

```

```

; APPLICANT: Cao, Yongwei

```

```

; APPLICANT: Edgerton, Michael D

```

```

; APPLICANT: Hinkle, Gregory J.

```

```

; APPLICANT: Kovalic, David K.

```

```

; APPLICANT: Liu, Jindong

```

```

; APPLICANT: Stein, Joshua

```

```

; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

```

```

; FILE REFERENCE: 38-10(52726)B

```

```

; CURRENT APPLICATION NUMBER: US/60/324,109

```

```

; NUMBER OF SEQ ID NOS: 33196

```

```

; SEQ ID NO 32880

```

```

; LENGTH: 504

```

```

; TYPE: PRT

```

```

; ORGANISM: Zea mays

```

```

; FEATURE:

```

```

US-60-324-109-32880

```

```

alignment_scores:

```

```

Quality: 1520.00
Ratio: 3.868
Percent Similarity: 80.698

```

```

Length: 487
Gaps: 4
Percent Identity: 61.602

```

```

alignment_block:

```

```

US-09-759-990-1 x US-60-324-109-32880 ..

```

```

Align seg 1/1 to: US-60-324-109-32880 from: 1 to: 504

```

```

19 ACTGGTGCCTTCATTCAGATACAGAAATGCCGACATCAACCTTCATGTTCT 68
|||||
27 TTTTSeTcIYArGTyrIleYrIleYsValIlyAspIleSerGIAlaAspH 43
CGCCCTTAAGAACTTACCTGCTGCTGAAGAAATGCCAGCTCTTATGG 118
|||||
43 eGIYArGIleuGIuIleIleuAlaGIuValGIleuMetProGIleuMet 60
TTCTTCGTAGCGTTATTCCTTCCTTAAGCATGAAGGTGTGTGAGATC 168
|||||
60 lAcYArGIAlaGIuPheGIYProSeTyrIlePheAlaGIYAlaArgIle 76
TCTGTGTTCCCTCCATGACATGACATGACAGACCGCTCCTCATGAGACT 218
|||||
77 SeGIYSeTleuHISmetThrIleGIuThrAlaValIleuIleGIuThrIle 93
CACAGCTTGTGCTGATGTGATGAGGCTTCTGCAACATCTTCTGTA 268
|||||
93 uThrAlaIleuGIYAlaGIuValArGIYrPCySeTyrCysAsnIlePheSeT 110
CACAAGATACAGCCGCTGCTGTATCGTGTGCGCCCAACAGCAGCACCA 318
|||||
110 hGIAlaSpIISAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 124
GAGAAGCCAGCCGCTATCCAGTTCCTGCTGGAAGGGCGAAACACATCCC 368
|||||
125 .....ValPheAlaIlePheGIYSeGIYGIuThrLeuGI 134
AGAACTACTGGAGAACATACGCGCTCTACATGCGCCAGATGCTCAG 418
|||||
134 uGIuYrIleTrrIlePcYrIleGIuArGIYSeTyrCysIleuAsPrrPheGIuAlaGIY 151

```

```

419 GCCCAGCAGAGGTTGTGATGATGATGGTGTGATGCTTACACTCTCATCTCC 468
|||||
151 lYProAspIleuIleValAlaSpAspIleYAlaSpAlaIleThrIleuIleHIS 167
AAGGCTTC.....GAATTCGAAACACCGGCTGCTTCCAGA 506
|||||
168 GIuGIYValIlyAlaGIuGIuAlaSpTyrGIuIlySeThrGIYIleProAs 184
GCCAACAAGACGTGACACTCGAATACCGCTGCTTGTCTACTCA 556
|||||
184 ProGIuSeTThrAspAsnAlaGIuPheIYIleValIleuThrIleIleA 201
AGCAGGCTTCAACCAAGACCAAGACACGCGCACAGTTCGTCGCCGC 606
|||||
201 rGAspGIYleuIlyAlaAspProIlyAlaSpTyrArGIYSeTyrGIuArG 217
ATGAACGCTGTTCCGAGAAGACACACACAGGTTCACCGCCTTACA 656
|||||
218 leuValGIYValIleSeTGIuGIuThrIleThrGIYValIlyArGIleuYrGI 234
GCTCGAGAAGAGGCGCAAACTCCTCTCCAGCCTACATCAACGACG 706
|||||
234 nMetGIuGIuTrGIYAlaIleuIleuPheProAlaIleAsnValAsnAsp 251
CTGTTCACAAAGTCCAAAGTTCGATTAACATCTACGCGTCCGCGACTCCT 756
|||||
251 eValIleThrIlySeTyrIlePheAspAsnIleuTyrGIYSeTyrGIuArG 267
ATCGATGATTCACACCGTCTCCGATTCATCATCAGAACTCGACCCAACTCGCT 806
|||||
268 ProAspGIYleuMetArGIAlaThrAspValMetIleAlaGIYAlaI 284
TCTGTCATGGGTACGGGATGTGCGCAAGGCTGCGCTCATTCCTCC 856
|||||
284 aValAlaIlySeTyrGIYAlaSpValGIYSeTyrGIYAlaIleuIle 301
GTGGCCAAAGCGCTCGCTTATCATCATCAGAACTCGACCCAACTCGCT 906
|||||
301 ySAlaIleGIYAlaArGIYValIleValIleThrGIuIleAspProIleCysAla 317
CTCCAGGCTGCTACGAGGCTACGAGGCTCCGCTCATGAGAAATCGT 956
|||||
318 leuGIuAlaIleuMetGIuGIuIleuIleuIleuProIleuGIuAlaI 334
CAAGATGTGCTATCTTCGTACATGCACAGAAATCGATATCATCT 1006
|||||
334 lSeTGIuAlaIleAspIlePheValIleThrIleThrIlyAsnIlyAspIleIleH 351
CTGTTCACATCATGCTCCAGATGAAGATTAAGCTATGTGCGGTATCATC 1056
|||||
351 eValAlaSpHISmetArGIYSeTyrIleYsAsnAlaIleValIlyAsnIle 367
GGCAGCTTCGATTAACGAATGATGATGACATGCGCTCATGAATATCCAG 1106
|||||
368 GIYHISpAsnGIuIleAspMetIleuGIYleuGIuThrIlyProGI 384
CATCAACACATCCCAATCAAGCCAGAAATAGCAGTGGGAATCCAG 1156
|||||
384 YValIlyArGIleThrIleIlySeTyrGIuIleThrAspArGIYrValPheProG 401
ATGGCCAC...GCTATCCTCTCTGCTGAGGCGCGCTTCTTAACCTT 1203
|||||
401 lUthrIleuIleGIYleIleValIleuAlaGIuGIYArGIleuMetAsnIleu 417
GGCTGCGTACAGGTACACCATCTTGTGTTATGCAATGTATATACAA 1253
|||||
418 GIYCysAlaIleThrGIYHISProSeTyrPheValMetSeTyrSePheThrAs 434
CCAGACACTGCTCAGCTCAGCTTACGAAAGAAAGA.....GGAATTC 1297
|||||
434 nGIuValIleIleAlaGIuIleuGIuIleuTrrIlyGIuIlySeTyrGIYSeT 451
TCGAGAAGAAGTTTACACACTCCGAAAGCATTCGATGAAGAAGTCCGT 1347

```



```

381 lYcYalAthrGlyHisProSerPheValMetSerCysSerPheThrAsn 397
1255 CAGACACTGCTCAGCTCGACCTCTACGAAAGAGA.....GGAATCT 1298
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 GluValAlaGlnLeuGlnLeuTyrPasnGlnLysAlaSerGlyLysTyr 414
1299 CAGAGAAAGGTTTACACACTTCCGAAAGCATCTCGAAGAGAGTCCGCTC 1348
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 rGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAlaL 431
1349 GCCTCACCTCGATCTCTCGATGTCACCTTACAAAGCTTACACAGAG 1398
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 euLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAsp 447
1399 CAGGCTGACTACATCAAGCTTCCAGTTGAGGCTCTTACAAGTCCGAGC 1448
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 GlnSerAspLysValSerIleProIleGluGlyProTyrLysProProH 464
1449 TTACGCTTAT 1458
|||||:|||||:
464 sTYrATgTYr 467

seq_name: /cgn2_6/ptodata/2/paa/US098_COMB pep:US-09-855-768-748

seq_documentation_block:
Sequence 748. Application US/09855768
GENERAL INFORMATION:
APPLICANT: Hauge, Brian M.
APPLICANT: Helne, Danielle L.
APPLICANT: Lin, Jie-Yi
APPLICANT: Parsons, Jeremy D.
APPLICANT: Wang, Ming Li
TITLE OF INVENTION: PLANT GENOME SEQUENCES AND USES THEREOF
FILE REFERENCE: 38-10(52047)B
CURRENT FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 09/760,427
PRIOR FILING DATE: 2001-01-13
NUMBER OF SEQ ID NOS: 1107
SEQ ID NO 748
LENGTH: 485
TYPE: PRT
ORGANISM: Glycine max
US-09-855-768-748

alignment_scores:
Quality: 1515.00 Length: 482
Ratio: 3.855 Gaps: 4
Percent Similarity: 81.535 Percent Identity: 61.826

alignment_block:
US-09-759-990-1 x US-09-855-768-748 ..
Align seg 1/1 to: US-09-855-768-748 from: 1 to: 485

34 GAGTACAGAAATTGGCGACATCACTCATGTCTCGCCGCTGAAGAACT 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTyrLysValLysAspLeuSerGlnAlaAspPheGlyArgLeuGluI 29
84 TACCTTGCTGAGAAAGAAATGCCAGGCTTATGTTCTTCGTCGAGAGCT 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 euLLeuAlaGluValGlnMetProGlyLeuMetAlaCysArgTyrThrGlu 46
134 ATTCGCGTTTAAAGCATTTGAAGGCTGTCAGAAATCTCTGCTCCCTCC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heGlyProSerGlnProPheLysGlyAlaArgIleThrGlySerLeuHis 62
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCAGACGCTTTGGTC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGlnIleThrLeuThrAlaLeuGly 79
234 TGATGTAGATGGGCTCTCTCGACATCTTCTTACCAAGATACAGCCG 283

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTyrPysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATTCGTTTGCGCCCAACAGGCACACACAGAACGCCACCGCT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 IaAlaAlaIleAlaArgAspSerAlaLa..... 105
334 ATCCAGCTCTCGCTGGAGGCGGAAACACTCCAGATACTGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTyrPylGlyThrLeuGlnGlyTyrTyrPry 120
384 CACATACCGGCTCTCATATGCCAGATGGTCAAGCCCAACAGAGTGG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sThrGluArgAlaLeuAspTyrGlyProGlyGlyProAspLeuIleV 137
434 TCGATGATGGGTGATGCTACATCTCCATCTCCAAAGGCTTCGAA... 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 AlAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153
481 .....TTCGAAACAGCCGCTGCTTCCAGAGCCCAACAGACGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GluGluLeuTyrGluLysThrGlyGluLeuProAspProAsnSerThrAs 170
522 CAACCTGGAATACCGCTCGCTTCTTGTCTACCTCAACAGAGTCTGAAC 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 PAsnAlaGlnPheGlnIleValLeuThrIleIleArgAspGlyLeuLysT 187
572 AAGACAGAACACACCTGGCCACAGTGTCTCCGCGCATGAACGCTTTTCC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 hAspProThrArgTyrArgLysMetLysGluArgLeuValGlyValSer 203
622 GAAGAGACACACAGGCTGCTCCAGGCTCTTACAGCTGAGAAAGAGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GluGluThrThrThrGlyValLysArgLeuTyrGlnMetGlnAlaAsnG 220
672 CAACCTGCTTCCAGCATCACTCACTCAACAGAGCTGTACAAAGTCCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 YThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSer 237
722 AGTTCGATTAACATCTAAGGCTGCGCCACTCCCTTATCGATGTATCAAC 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 ySPheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGTTCGATGATGATGATGCGGCGGCAAGACAGCTGTCAATGGGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValAlaAlaGlyTyr 270
822 CGGCGATGTGCGCAAGGCTGCGCTCATATCCCTCGTGCCCAAGGCTGC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGlyAspValGlyLysGlyCysAlaAlaAlaMetLysGlnAlaGlyAla 287
872 GCGTTATCATCACAGACTCGAACCCAAATCTGCGCTTCCAGGCTGGCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rGValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAAGCTACAGAGTCCGCGCCGATCGAGAGATGTCAGATGTCATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GluGlyLeuGlnValLeuThrLeuGlnLysValValSerGlnAlaAsp 320
972 CTTCGTTATCATCACAGAAACTGCGATATCATCTCTGTGACATGATGG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValAspHisMeta 337
1022 CCCAGATGAAGATAGCGTATGTGGGTAACTCGGCACTCGCCACTTGATAAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 rGlySerLysAsnAlaIleValAlaCysAsnIleGlyHisPheAspAsn 353
1072 GAAATGTATACAGATGGCTCATGAAATATACAGGATCAAGCATGCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GluIleAspMetLeuGlyLeuGlnLysValAsnTyrProGlyValLysArgIle 370
1122 AATCAAGCCAGATACAGATGTGGGAATCCCGAGATGGCCAC...GCTA 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

370 rtleysproglinThrAspArgTrpValIlePheProglinThrAsnThrGlyI 387
1169 TCCTCCCTTCCTGCTGAGGCGCCCTTCTTAACCTGGCTGCGTACAGGT 1218
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 leileValleuIaIaIaGlyArgIleuMetAsnIleuGlyCysAlaThrGly 403
1219 CACCATTCTTGGTATGTCATATGTCATTCACAACACAGACATCGCTCA 1268
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 HisProSerPheValMetSerCysSerPheThrAsnIleValIleAlaGI 420
1269 GCTGACCTCTACAGAAAGAGA.....GGAATCTGAGAAAGAGTTT 1312
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 nleuGluIeuTrpLysGluLysSerThrGlyLysTyrGluLysValIT 437
1313 ACACACTTCGAGACATCTGATGAGAAAGTGGCTGCTCCACCTCGGA 1362
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 yValIleuProLysHisIleuAspGIuLysValAlaIaIaIeunHisIleuGI 453
1363 TCTCTCGATGTCACCTTACAAAGCTTACACAGACAGAGCTGACTACAT 1412
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 LysIleuGIuLysAlaLysIleuThrGluIleuSerLysSerGIuLysAspTyrII 470
1413 CAACGTTCCAGTTGAGGCTCTTACACAGTGTGATGCTTACCGTTAT 1458
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 eSerValIProValGIuGIuLysProTyrLysProAlaHisTyrArgTyr 485

```

seq_name: /cgn2.6/plodata/2/paa/US60.COMB.pep:US-60-324-109-30679

seq_documentation block:

```

; Sequence 30679, Application US/60324109
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Steiu, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)B
; CURRENT APPLICATION NUMBER: US/60/324.109
; NUMBER OF SEQ ID NOS: 33196
; SEQ ID NO 30679
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
US-60-324-109-30679

```

alignment_scores:

```

      Quality: 1515.00      Length: 482
      Ratio: 3.855          Gaps: 4
      Percent Similarity: 81.535      Percent Identity: 61.826

```

alignment_block:

US-09-759-990-1 x US-60-324-109-30679 ..

Align seq 1/1 to: US-60-324-109-30679 from: 1 to: 485

```

34 GAGTACAGAAATTCGCGACATCCATTCCTGCGCGCTAGAGAACT 83
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTrpLysValLysAspLysSerGIuLysAspPheGIuArgIleuGIuII 29
84 TACCTTGTCTGAGAGAAATGCGACGCTTATGTTCTTCGTGAGCGTT 133
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 eGIuIeuIaGIuValGIuMetProGIuLysMetAlaCysArgThrGluP 46
134 ATTCGCTCTTAAGCATTTGAAGGGTGTGAGAAATCTGTGCTCCCTCAC 183
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heGIuProSerGIuProPheLysGIuLysAlaArgIleThrGIuSerLeuHis 62
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCAGACTCTTGTCGC 233

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGIuThrAlaValIleuIleGIuThrLeuThrAlaIeGIuAl 79
234 TGATGCGAGATGGGCTTCGCAACATCTCTCTACACAAATACAGCCG 283
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGIuValArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATCTGTTGCGGCCAACAGGCACACAGACAGACCCGCT 333
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 IaAlaIaIeAlaIaArgAspSerAlaIa..... 105
334 ATCCAGCTTTCGCTGGAAGGCGGAACACTCCAGAAATCTGGAGAA 383
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGIuGIuThrIleuGIuIleuTyrTrpC 120
384 CACATACCGCGCTTCACATGGGCAGATGGTCAGAGGCCACAGCGTTG 433
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sThrGIuIaArgAlaIeAspTrpGIuProGIuGIuGIuProAsnIleu 137
434 TCGATGATGGTGGTGTGATGTACACTCCTCATCTCCAGGCGCTCGAA... 480
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 aIAspAspGIuLysAlaIaThrIleuIleuHisGIuGIuValLysAla 153
481 .....TTGGAACAGCCGCTGCTTTCAGAGCCACAGAAAGCTGA 521
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GIuGIuLysTrpGIuLysThrGIuLysProAsnSerThrAs 170
522 CAACCTCGAAATACCGCTGCTTCTTACACAGAGGCTCTCAACC 571
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 PAsnIaGIuPheGIuIleValIleuThrIleIaArgAspGIuLysTr 187
572 AAGACAGACACACCTGGCACACAGTTGCTGCGGCAGATGAGCGTGTCC 621
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 hrAspProThrArgTyrArgLysMetLysGIuArgIleuValGIuValSer 203
622 GAAGAGACACACAGAGTGTCCACCGCTTACACGCTCGAAGAGAGG 671
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GIuGIuThrThrThrGIuLysArgLysTrpIleuMetGIuAlaAsnGI 220
672 CAACCTCTCTTCCAGCATCAAGCTCAAGCAGAGCGTGTCAAAAGTCA 721
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 yThrIleuIeuPheProAlaIleAsnValAsnAspSerValThrLysSer 237
722 AGTTGATATACATAGGCTGCGCGCATCTCCCTTATGATGATGATCAAC 771
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 ySPheAspAsnIleuTyrGIuArgHisSerIleuProAspGIuLysMet 253
772 CGTCTTCCGATGTCTATGATCGCGCGCAACAGACTCTCGTATGGGTTA 821
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ArgAlaThrAspValMetIleAlaGIuLysValAlaValAlaIaGIuTy 270
822 CGGCGATGTCGCGAAGGCTGCGCTCAATCCCTCGGCGCCAGCGCTC 871
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGIuAspValGIuLysGIuCysAlaIaAlaIaMetLysGIuIaGIuAla 287
872 GCGTTATCATCAAGAACTGACCAATCTGCGCTTCACAGCTGCCATG 921
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rGIuAlIleValThrGIuIleAspProIleCysAlaIeGIuAlaIeMet 303
922 GAAGCTACCAAGCTCCGCGCATCGAGAAAGTGTCAAGATGTCGATAT 971
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GIuGIuLysGIuValIleuThrIleuGIuAspValAlaIaSerGIuAlaAspII 320
972 CTTCGTACATGCACAGAAACTGCGCATTCATCTGTCATGATGATGG 1021
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrThrGIuAsnLysAspIleIleMetValAspHisMet 337
1022 CCCAGATGAAGATAGCTATGTGCGTACATCGGCGCACTTGATTAAC 1071
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 rGIuMetLysAsnAlaIleValLysAsnIleGIuHisPheAspHis 353
1072 GAAATGTATACAGATGCGCTCATGAATACCCAGGCTACAGCATCC 1121
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

354 GluIleaspMetLeuGlyLeuGluIuAsnTyrProGlyValIlysArgIleth 370
1122 AATCAACGCGAATACGACATGTGGGAATTCAGATGGCCAC...GCTA 1168
|||||
370 rIleValProGlnThrAspArgTTPValPheProGlnThrAsnThrGly 387
1169 TCCTCTCTTCCTGAGGCGCCCTTCCTAACCTGGCTGCGCTACAGGT 1218
|||||
387 leIleValIleuAlaGluGlyArgIleuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCAATCTTTCGTTATGTCATTCATTCACAAACACAGACTCGCTCA 1268
|||||
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420
1269 GCTGACCTTCAGAAAGAGA....GGAATCTCGAGAAAGAGTTT 1312
|||||
420 nIleuGlnThrPlySgluYsSerThrGlySlyrGluYsIlyValI 437
1313 ACACACTTCGAGACATCTCGATGAAGAAGTCGCTCGCTCCACTCGGA 1362
|||||
437 yIleValProGlyHisIleuAspGluYsValIAlaIleuHisLeuGly 453
1363 TCCTCGATGTCACACTTACAAAGCTTACAGAGAGAGAGCTGATACAT 1412
|||||
454 LysLeuGlyAlaLysLeuThrGlnLeuSerLysSerGlnAlaAspTyr 1470
1413 CAACGTTCCAGTGTGAGGCTCTTACAGCTGATGCTTACCGCTTAT 1458
|||||
470 eSerValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485

```

```

seq_name: /cgn2.6/prodata/2/paa/US097_COMB.pep:US-09-708-427-22306

```

```

seq_documentation_block:
: Sequence 22306, Application US/09708427
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 22306
: LENGTH: 450
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..450
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..450
: OTHER INFORMATION: Ceres Seq. ID 1840724
: US-09-708-427-22306

```

```

alignment_scores:
  Quality: 1486.00      Length: 459
  Ratio: 3.963          Gaps: 4
  Percent Similarity: 81.699  Percent Identity: 63.181

```

```

alignment_block:
US-09-759-990-1 x US-09-708-427-22306 ..

```

```

Align seg 1/1 to: US-09-708-427-22306 from: 1 to: 450

```

```

103 ATGCCAGCTTATGAGTCTTCGAGCGTATTCGCTTAAAGCAT 152
|||||
1 MetProGlyLeuMetAlaCysArgThrGluPheGlyProSerGlnPro 17
153 GAAGGCTCAGAAATCTGCTGCTCCCTCCACATGACAGTCCAGAGCCG 202
|||||

```

```

17 eLysGlyAlaArgIleThrGlySerLeuHisMetThrIleGlnThrAlaVal 34
203 TCCCATATCGAGACACTCACAGCTCTGTGCTGATGTCAGATGGGCTTCC 252
|||||
34 aIleuIleGluThrLeuThrIleAlaLeuGlyAlaGluValAlaGlyTyrCys 50
253 TGCACATCTTCTTACACAGATACAGCCGCTGCTGTATGCTGTGCGG 302
|||||
51 CysAsnIlePheSerThrGlnAspHisAlaIleAlaIleAlaIleAlaArgAs 67
303 CCCAACAGGACACACAGAGAGAGCCGCTATCCAGTCTTGCCCTGGA 352
|||||
67 pSerIleAla.....ValIleAlaThrPl 75
353 AGGCGAAACACTCCAGATATCTGGAGAACACATACCCGCTGCTACA 402
|||||
75 ySglGlyThrLeuGlnGluTyrTyrTyrCysThrGluAlaAlaLeuAsp 91
403 TGGCCAGATGTCAGAGCCACAGAGGTTGTCGATGATGCTGATGC 452
|||||
92 TrpGlyProGlyGlyProAspLeuIleValAspAspGlyAspAl 108
453 TACACTCTCATCTCCAGAGGCTTCGAA.....TTCGAACAG 490
|||||
108 aThrLeuLeuIleHisGluGlyValIlyAlaGluGluIlePheGluTyr 125
491 CCGGTGCTTTCAGAGCCACAGAGAGTGCACACCTCGAATACCGCTGC 540
|||||
125 hGlyGlnValProAspProThrSerThrAspAsnProGluPheGlnIle 141
541 GTTCTTGTACACTCAAGAGGCTTTCACCAAGACACAGAACCACTGGCA 590
|||||
142 ValLeuSerIleIleLysGluGlyLeuGlnValAspProLysTyrYr 158
591 CACAGTGTGCGGCGATGAGAGGTTTCGGAAGAGACACACAGCAGTGC 640
|||||
158 sLysMetLysGluArgLeuValGlyValSerGluGluThrThrGly 175
641 TCCACCGCTCTTACAGCTCGAGAGAGAGGCAAACTCTTCTCCAGCC 690
|||||
175 aLysArgLeuThrGlnMetGlnGlnAsnGlyThrLeuLeuPheProAla 191
691 ATCAACGTACAGAGCGCTTACAAAGTCCAGTTCATACATCTACGG 740
|||||
192 IleAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTyr 208
741 CTGCGGCCACTCCCTATGATGATACCAACCGGCTCGATGATCAATGA 790
|||||
208 yCysArgHisSerLeuProAspLysLeuMetArgAlaThrAspAlaMet 225
791 TCGGGGGCAAGACAGCTCTGTCATGGGTTACGGCGATGTCGGCAAGGC 840
|||||
225 leAlaGlyLysValAlaValIleCysGlyTyrGlyAspValGlyGly 241
841 TGCAGTCAATCCCTCCGTCGAGGCGCTCGGCTTATCATACAGAACT 890
|||||
242 CysAlaAlaAlaMetLysThrAlaGlyAlaArgValIleValThrGlu 258
891 CGACCAATCTGCGCTCCAGGCTGCGCATGAGAGGCTACAGAGTCCGCC 940
|||||
258 eAspProIleCysAlaLeuGlnAlaLeuMetGluGlyLeuGlnValLeu 275
941 GCATCGAGAGAGTGTCAAGGATGTCGATATCTTCTGTTACATGCAAGGA 990
|||||
275 hIleGluIuAspValValSerGluAlaAspIlePheValThrThrThrGly 291
991 AACTGCGATATCATCTCTGTGACATGATGCGCCAGATGAAGATTAAGGC 1040
|||||
292 AsnLysAspIleIleMetValAspHisMetArgLysMetLysAsnAspAl 308
1041 TATTGCGGTACATCGGCGACTTGCATACGAATGATGATACAGATGGCC 1090
|||||
308 aIleValCysAsnIleGlyHisPheAspAsnGluIleAspMetLeuGlyL 325

```


[illegible]

seq_name: /cgn2_6/ptodata/2/paa/US096_COMB.pep:US-09-620-394B-3583

seq_documentation_block:

; Sequence 3583, Application US/09620394B

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai

APPLICANT: BROVER, Vyacheslav

TITLE OF INVENTION: Sequence-Dependent Control of Transcription

TITLE OF INVENTION: Thereby
; FILE REFERENCE: 3750-1067D

FILE REFERENCE: 2750-1067p

```

; CURRENT APPLICATION NUMBER: US/
CURRENT FILING DATE: 2000-07-2

```

NUMBER OF SEQ ID NOS: 9131

NUMBER OF SEQ ID NOS: 9151
SEQ ID NO 3583

SEQ ID NO 5283
LENGTH: 467

TYPE: PRT

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1..467

OTHER INFORMATION: Xaa 1s any

```
NAME/KEY: misc_feature
```

LOCATION: 1..467

OTHER INFORMATION: Ceres Seq.

alignment_scores:

Quality: 1457.00

Ratio: 3.865

Percent Similarity: 80.901

—

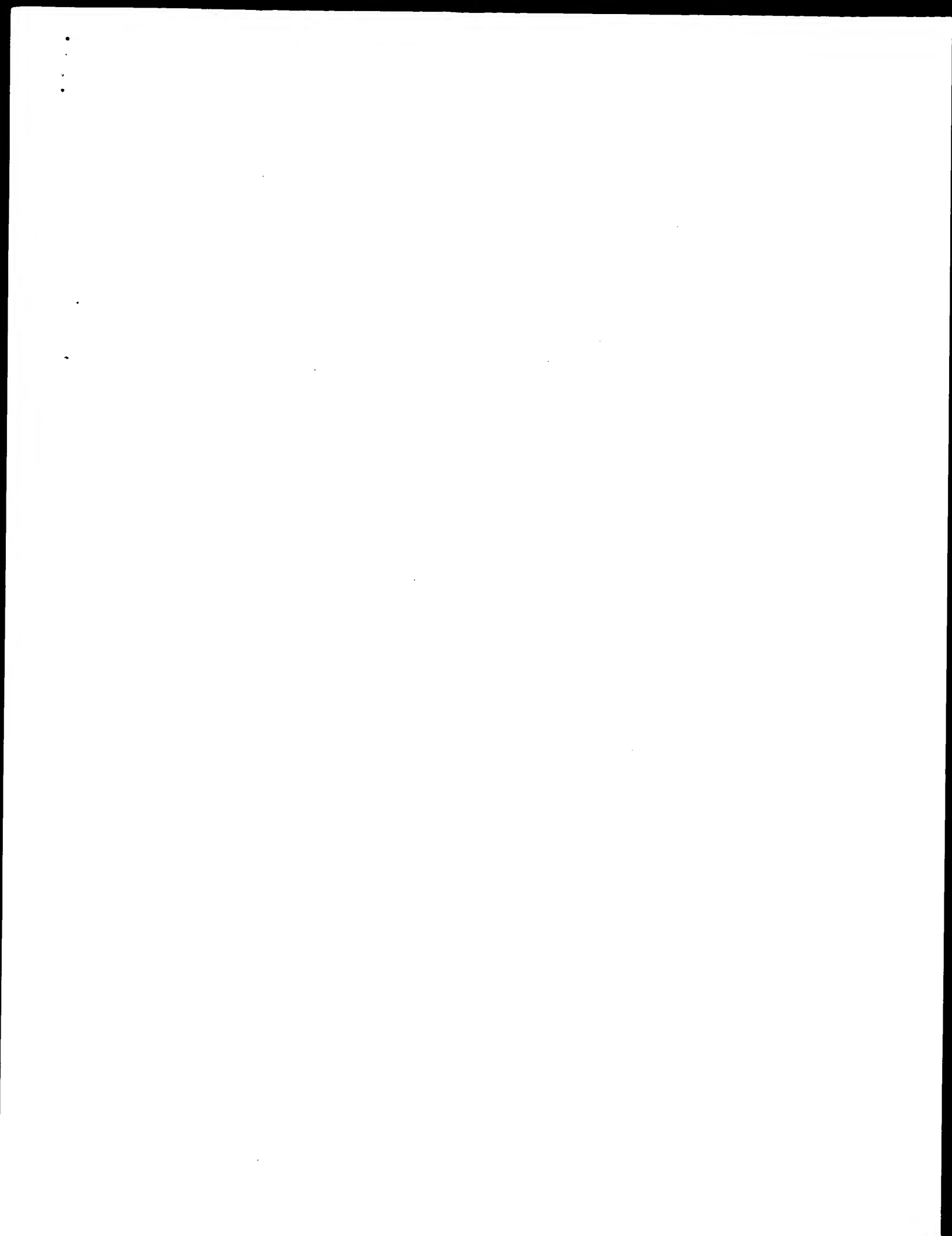
```
alignment_block:
```

US-09-759-990-1 x US-09-620-394B-3583

TC-00-630-30AD-3583 FROM: 1 TO: 467

[illegible]

```
||||| ||||| ||||| ::|||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGlnAspValVa 315
957 CAAGAGTGTGATATCTTGTACATGCACAGAAACTGGCATATCATCT 1006
||||| ||||| ||||| |||||
315 IserGlnAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
1007 CTGTTGACATGATGGCCGATGAGGATTAAGGCTATTGCGTAACATC 1056
||||| ||| ::||| ||||| |||||
332 etValAspHisMetLargLysMetLysAsnAlaIleValCysAsnIle 348
1057 GGCCACTTGCATACGAAATTTGATACAGATGCCCTCATGAAATACCCAG 1106
||||| ||||| ||||| |||||
349 GlyHisPheAspAsnGlnIleAspMetLeuGlyLeuGlnThrTyrProG 365
1107 CATCAAGCACATCCCAATCAAGCCAGAAATACGACATGTGGGAATCCGAG 1156
::||| ||||| ||||| |||||
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrPvalPheProG 382
1157 ATGGCCAGCT...ATCTTCCTTCTTGTGAGGCGCGCTTCTTAACCTT 1203
:: ||| ||||| ||||| ||||| |||||
382 IuThrLysAlaGlyIleIleValLeuAlaGlnGlyArgLeuMetAsnLeu 398
1204 GGCTGGCTACAGGTCACCCATCTTTCGTTATGTCATGTCAATGCACAA 1253
||||| ||||| ||||| ||||| |||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CGAGACACGCGCTGAGCTGACCTCTAGAAAGAGA.....GGAAATC 1297
||||| ||||| ||||| ||||| |||||
415 ngInValIleAlaGlnLeuGlnLeuTyrPasnGlnLysAlaSerGlyLysT 432
1298 TCGAGAAAGAAGTTTACACACTTCGAGACATCTCGATGAGAGAGTCGCT 1347
||||| ||||| ||||| ||||| |||||
432 yTGlnLysLysValTyrValLeuProLysHisLeuAspGlnLysValAla 448
1348 CGCCTTCACCTCGGATCTCTGATGTCACCTTACAAAGCTTACACAG 1395
||||| ||||| ||||| ||||| |||||
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLys 464
```



OM of: US-09-759-990-1 to: Pending_Patents_AA_New.* out_format : pfs
Date: Aug 17, 2002 1:44 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame+nt2p.model -DEV=rlp  
-Q=/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 2534.00 3228.77 4.7e-172  
-DB=pending_Patents_AA_New -QEMT=fastan -SUFFIX=nt2p.rapn  
-GAPDP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000  
-LOOPEXT=0.000 -OGAPDP=4.500 -OGAPEXT=0.050 -XGAPDP=10.000  
-XGAPEXT=0.500 -FGAPDP=6.000 -FGAPEXT=7.000 -YGAPEXT=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1  
-MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200  
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US0975990_@cgn1.193 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1
```

Search information block:

Query: US-09-759-990-1
Query length: 1461
Database: Pending_Patents_AA_New.*
Database sequences: 710927
Database length: 235927762
Search time (sec): 174.720000

score list:

Sequence	Strd orig	Zscore	EScore	len	Documentation	
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 2534.00 3228.77 4.7e-172						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1724.00 2192.81 2.4e-114						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1645.00 2091.56 1.0e-108						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1588.00 2018.35 5.7e-102						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1550.50 1970.47 5.7e-102						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1544.00 1962.36 1.6e-101						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1537.00 1953.41 5.2e-101						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1536.00 1952.13 6.1e-101						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1533.00 1948.29 1.0e-100						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1524.00 1936.78 4.4e-100						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1520.00 1931.66 8.4e-100						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1518.00 1929.49 1.2e-99						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1518.00 1929.49 1.2e-99						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1512.00 1921.43 3.1e-99						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1509.00 1917.59 5.1e-99						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1506.00 1913.75 8.4e-99						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1503.00 1910.30 1.4e-98						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1500.00 1906.42 2.2e-98						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1499.00 1904.80 2.6e-98						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1499.00 1904.80 2.6e-98						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1494.00 1898.79 5.9e-98						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1475.00 1874.86 1.3e-96						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1474.00 1873.18 1.6e-96						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1468.00 1865.51 4.2e-96						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1442.00 1832.40 3.0e-94						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1434.00 1822.51 1.1e-93						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1415.00 1797.82 2.5e-92						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1411.00 1792.66 4.9e-92						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1406.00 1786.69 1.1e-91						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1402.00 1781.23 2.1e-91						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1395.00 1772.32 6.7e-91						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1382.00 1755.65 5.6e-90						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1382.00 1755.65 5.6e-90						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1374.00 1745.37 2.1e-89						
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1374.00 1745.37 2.1e-89						

```
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1369.00 1738.98 4  
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1364.00 1732.58 1  
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1363.00 1731.26 1  
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1359.00 1726.18 2  
/cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231 + 1356.50 1722.92 3
```

seq_name: /cgn2_6/pdata/1/paa/US09_NEW_COMB.pdp:US-09-791-537-49231

seq_documentation_block:

```
; Sequence 49231, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMI  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 49231  
; LENGTH: 486  
; TYPE: PRT  
; ORGANISM: Trichomonas vaginalis  
US-09-791-537-49231
```

alignment_scores:

Quality	2534.00	Length:	486
Ratio:	5.225	Gaps:	0
Percent Similarity:	99.794	Percent Identity:	98.971

alignment_block:

US-09-759-990-1 x US-09-791-537-49231 ..

Align seg 1/1 to: US-09-791-537-49231 from: 1 to: 486

```
1 ATGGCTTGAATCAATCACTACTGCTGCTCCATTCAGTACGAAATTCGCGA 50  
1 MetAlacysylsSerProAlaGlyAlaProPheGluThrArgIleAlaAs 17  
51 CATCAACCTCATGTTCTGCGCGTAAGGAACTTACCTTGTCAGAGAAG 100  
17 pIleAsnLeuHisValLeuGlyArgIleGluThrLeuAlaGluIulysG 34  
101 AAATGCCAGGCTATGCTGCTGCGGCTATTCCTGCTTACGCCA 150  
34 LunetProGlyLeuMetValLeuArgGluArgTyrSerAlaSerLysPro 50  
151 TTGAAGGCTGTCAGAACTCTGCTGCTCCATTCAGTACGAGCAGC 200  
51 LeuIleGlyValAlaGlyIleSerGlySerLeuHisMetThrValGlnThrAl 67  
201 CGTCCATATGAGACACTCAAGCTCTGCTGCTGCTGCTGCTGCTGCT 250  
67 aValLeuIleGluThrLeuThrAlaLeuGlyAlaAspValArgTyrPAlas 84  
251 CTTGCAACATCTCTTACACAGATACAGCGGCTGCTGCTGCTGCTGCT 300  
84 eCysAsnIlePheSerThrGlnAspThrAlaIleAlaIleValVal 100  
301 GGGCCCAAGGACGACACAGAGAACCGGATCCAGTTCGCTGCTGCT 350  
101 GlyProThrIleThrProGluThrProAlaGlyIleProValPheAlaIrr 117  
351 GAAGGCGAAGACCTCCAGATATCGGAGAACACATACCGCGCTCTCA 400  
117 pIySglGluThrLeuProGluThrProGluThrProGluThrProGluThr 134  
401 CATGGCAGATGCTCAAGGCGACAGCAGAGGTTGATGATGATGATGAT 450  
134 hTrpProAspGlyGlnGlyProGlnGlnValAlaAspAspGlyGlyAsp 150
```



```

109 ValProValPheAlaTrpLysGlyIuThrLeuGluIuTrpTrpCys 125
384 CACATACCGGCTCTCACATGGCCAGATGGTCA.....GGCCACAGC 427
125 sthrgluGlnAlaLeuThrTrpProAsnThrProThrglyGlyProAsn 142
428 AGGTTTCGATGATGGTGTGATGCTACACTCTCATCTCCAGAGGCGCTC 477
142 eLLeuAspArgLysGlyAspAlaThrLeuValHisLysGlyVal 158
478 GAATTTCGAACACCGGCTGCTTCCAGACGCCAACAGAGAGTGAACACT 527
159 GluTrpGluLysAspGlyValProSerValAspThrAlaGluSerAs 175
528 CCAATTCGCGTGGCTCTTCTACACTCAACAGAGTCTTCAACACAGACA 577
175 pGluHisArgValIleLeuGluIuLeuThrArgThrValGlyGluSerP 192
578 AGAACCACTGGCACACAGTTGCTGGCGGCATGAACGCTGTTCCGACAG 627
192 roGlnLysTrpThrGlnIleuAlaSerGluIleArgGlyValThrGlu 208
628 AACACACAGGCTGTCCACGCGCTCTACAGCTCGAAGAGGCGCAACT 677
209 ThrThrThrGlyValHisArgLeuTrpGluMetHisArgAspGlyThr 225
678 CCTCTCCAGACCATCAACGCTCAACGAGCGGTGTACAAAGTCAAGTTCG 727
225 uLeuPheProAlaIleAsnValAsnAspAlaValThrLysSerLysPhe 242
728 ATPACATCTACGGCTGGCGGCACTCCCTTATGATGGTATCAACCGTCT 777
242 sPAsnLysTrpGlyCysArgHisSerLeuIleAspGlyIleAsnArgAla 258
778 TCGAGTGTATGATCGGCGGACAGACAGCTCTGTCATGGTATTCAGGGA 827
259 ThrAspValLeuIleGlyGlyLysThrAlaValAlaCysGlyTrpGly 275
828 TGTGCGCAAGGCTGGCTCATCTCCCTCGGCAAGCGCTCGCTTA 877
275 pValGlyLysGlyCysAlaGluSerLeuArgGlyGlnGlyAlaArgVal 292
878 TCATCAGACACTGACCCCAATCTGCGCTTCCAGCGCTCCATGGAAGC 927
292 LeLeuThrGluIleAspProIleCysAlaLeuGlnAlaIleMetAsp 308
928 TACAGGTCGCGCGCATCGAGAGAGATGTCAGATGATGATCTGCTGT 977
309 pHeGlnValThrThrLeuAspGluValValAspLysAlaAspIlePhe 325
978 TACATGCACAGAACTGCGATATCATCTGTGTGACATGATGCGCCAGA 1027
325 LThrThrThrGlyAsnLysAspIleIleMetAlaLysAspMetAlaLys 342
1028 TGAAGGATAGGCTATTTGCGTAACTGCGCACTTCGATACCAACTT 1077
342 eLysHisGlnAlaIleValGlyLysIleGlyHisPheAspAsnGlu 358
1078 GATACGATGCGCTCATGAATACCGAGCATCAAGCACATCCCAACTCA 1127
359 AspMetAlaGlyLeuAlaGlnThrProGlyIleValLysAspGluVal 375
1128 GCCAGAAATACGACATGTGGAAATCCAGATGGCCAGCTATCTCTTC 1177
375 sProGlnValHisThrThrTrpTrpArgAspGlyLysValLeuIleVal 392
1178 TTGCTAGAGGCGGCTTAACTTGGCTCGCTACAGCTCACCCCACT 1227
392 euserGluGlyArgLeuLeuAsnGluLysAlaThrGlyHisProSer 408
1228 TTGCTATGTCATGTCTACAAACACAGACTCGCTAGCTGACTGACT 1277
409 PheValMetSerAsnSerPheAlaAspGlnThrLeuAlaGlnIleGlu 425

```

```

1278 CTACGAAAGAGAGAAATCTCGAAGAGAGGTTTACACACTTCCGAGC 1327
425 uPheThrLysProAspGlyTrpProThrAspAlaGlyValIleuProLysH 442
1328 ATCTGATGAGAAGTGGCTGGCTCCACACTTCGATCTCTCGATGTCAC 1377
442 ILeuAspGlyLysValAlaArgLeuHisIleAspSerLeuGlyValLys 458
1378 CTTCGAAAGCTTACACAGACGAGTGCATCATCAACGTTCCACTTGA 1427
459 LeuThrThrLeuArgProGluGlnAlaAspTrpIleGlyValLysVal 475
1428 GCGTCTTACAACTGTGATGCTTACCGTTAT 1458
475 uGlyProLysAlaAspHisLysTrpArgTrp 485

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-540-209B-7902

seq_documentation_block:
; Sequence 7902, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOID
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7902
; LENGTH: 499
; TYPE: PRN
; ORGANISM: B. fragilis
US-09-540-209B-7902

alignment_scores:
Quality: 1588.00 Length: 476
Ratio: 3.950 Gaps: 4
Percent Similarity: 84.454 Percent Identity: 64.916

alignment_block:
US-09-759-990-1 x US-09-540-209B-7902 ..

Align seq 1/1 to: US-09-540-209B-7902 from: 1 to: 499

37 TACAGAAATGCGGACATCAACCTTCATGCTGCGCGTAAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
38 TyrLysValAlaAspIleThrLeuAlaAspPheGlyArgGlyGluIleAs 54
87 CTTTGTGAGAGAGAAATGCGAGTCTTATGTTGTTCTTCTGTAAGCTTAT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 pLeuAlaGluLysGluMetProGlyLeuMetAlaLeuAlaGlyLysGly 71
137 CCGCTTAAACCAATTGAAGGCTGCAGAACTCGTTCGCTCCGTCACAG 186
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 LysLysLeuLysProLeuLysGlyAlaArgIleMetCysSerLeuHisMet 87
187 ACAATCCAGACAGCGCTCTCATCGAGACACTCACAGCTTCTGGTGGTCA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 ThrIleGlnThrAlaValIleLeuIleGluThrLeuValAlaLeuGlyAla 104
237 TGTGATGATGGCTTCTGCAACATCTTCTTACACAGATACAGCCGGTGG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
104 uValAlaTrpLysSerGlyAsnIleLysSerThrGlnAspHisAlaAla 121
287 CTGCTATGCTGTGTGGCCCAACAGGCACACAGAAAGCAAGCCGCTATC 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AlaAlaIleAlaAla.....serGlyVal 128
337 CCACTCTTGGCTGGAAGGCGAAACACTCCAGAAATACTAGAGGAACAC 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 AlaValPheAlaTrpLysGlyGluThrLeuAlaAspTrpTrpTrpCysH 145

```



```

120 ValProValPheAlaTrpLysGluTrpLeuGluTrpTrpAl 136
136 CACATACGCGCTTCACATGGCCAGATGGTCAAGGCCA...CAGCAGG 430
136 AlaAlaGluMetLeuThrTrpProAspProAspLysProAlaAsnMetI 153
431 TTGCGATGATGGTGGTGTGTACACTCCCTCATCTCCAAAGGGCTTGAA 480
153 IeLeuAspAspLysGlyAspAlaThrMetLeuValLeuArgGlyMetGln 169
481 TTCGAAACAGCGGCTGCTTCAGAGCCACAGAACTCAACAACCTCGA 530
170 TyrGluLysAlaGlyValAlaProProAlaGluGluAspAspProAlaG1 186
531 ATACGCTGCGCTTCTGTACACTCAAGAGCTTCACACCAAGACAGAGA 580
186 uTrpLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCACTGGCACACAGTGTGCTGCGCGCATGAAAGGTGTTCGAGAGACA 630
631 ACAACAGGTGTCCACCGCTCTACAGCTCGAGAGAGAGGCAACTCTCT 680
220 ThrThrGlyValLeuArgLeuTrpGlnPheAlaAlaAlaLysAspLeuAl 236
681 CTTCGCCACCATCAACGTCACAGCAGCTGTTCACAAAGTCCAGTTCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGGCTGCGCGCACCTCCCTTATGAGGATCAACCGTGTCTCC 780
253 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATGCGGCGCAAGACAGCTTCGTCATGGGTACAGCGCATGT 830
270 AspAlaLeuIleGlyLysLysValLeuIleCysGlyTyrLysAspVal 286
831 CGGCAAGGGCTGCGCTCAATCCCTCCGTCGCAAGCGCTGCGGTATCA 880
286 IGlYysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTGCACCCATCTGCGCTTCACAGGCTGCCATGAGAGCTTAC 930
303 aThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGTCCCGCGCATGAGAACTGCTCAAGAGATGATGATCTTCTCTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValAlaThr 336
981 ATGCACAGAACTGCGATATCATCTCTGTGATCATGATGAGCCAGATGA 1030
336 rAlaThrLysAsnLysAspIleIleMetLeuGluHisIleLysAlaMetL 353
1031 AGATTAAGCTATTGTGCGTAACTGCGCACTTCGATTAAGAAATTGAT 1080
353 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGCGCTCATGAATACCCAGCATCAAGCATCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysSpr 385
1131 AGAATACACATGTGGAAATCCAGAT...GGCCACGATATCTCTCTG 1177
385 oGlnValAspLeuTrpThrPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGCTGAGGCGCGCTTAACTGTGCGCTACAGCTACAGCTACCCATCT 1227
402 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTCGTATGTCATGTCAATCAACAGACATCGCTCAGCTCAGCT 1277

```

```

419 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTACGAAAGAGAGAGAAATCTCGAGAAAGAGCTTACACACTCCGAGAC 1327
435 uTrpThrLysAsnAspLysValLysArgLysValAlaLysArgLeuProLysH 452
1328 ATCTCGATGAAGAGATGCTGCGCTCCACCTCGGATCTCTCGANGTCCAC 1377
452 ILeuAspGluLysValAlaArgIleHisValGluAlaLeuGlyLysHis 468
1378 CTTCACAAAGCTTACACAGAGAGAGCTGACACTACATCAACGTTCCAGTTGA 1427
469 LeuThrLysLeuThrLysGluGlnAlaGluLeuTrpLeuGlyValAspValG1 485
1428 GGGTCTTACAGTGTGATGCTTACCGTTAT 1458
485 uGlyProTyrLysProAspHisTyrAlaGlyTyr 495

```

seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-37085

seq_documentation_block:

Sequence 37085, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

NUMBER OF FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 37085

LENGTH: 485

TYPE: PRT

ORGANISM: Petroselinum crispum

US-09-791-537-37085

alignment_scores:

Quality: 1544.00 Length: 488

Ratio: 3.879 Gaps: 6

Percent Similarity: 81.557 Percent Identity: 63.115

alignment_block:

US-09-759-990-1 x US-09-791-537-37085 ..

Align seg 1/1 to: US-09-791-537-37085 from: 1 to: 485

```

19 ACTGGTGTCCATTCAGATGACAAATGGCGACATCAACCTCATGTCT 68
8 ThrAlaAlaGlyArgGluTrpLysValLysAspMetSerLeuAlaAspH 24
69 CGGCGCTAAGAACTTACCCTTCTGTAAGAAAGAAATCCAGCTTATAG 118
24 eGlyArgLeuGluLeuGluLeuAlaGluValAlaGluMetProGlyLeuMetS 41
119 TTCTTCGTAGAGCTTATCCGCTTCTTAACCATTAAGAGGGTGTGAGATC 168
41 eCysArgThrGlnPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169 TCTGGTTCCTCCACATGACAGTGCACAGAGCCGCTCATGAGAGACT 218
58 ThrGlySerLeuHisMetThrIleGlnThrGlyAlaLeuIleGluThrIle 74
219 CACAGCTTGTGTGTGATGTCAGATGGCTCTCGACAACTTCTCTTA 268
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 318
269 CACAGATACAGCGCGCTGCTGCTATGCTGTTCGGCCCAACAGGACACA 318
91 hGlnAspHisAlaAlaAlaIleAla..... 100

```

319 GAGAACCCACGCGGTATCCAGACTCTTGCGCGGAAGGGCGAAACACTGCC 368
101 ArgAspSerCysAlaVal.PheMetAlaTrpLysGlyIleThrLeuGI 115
369 AGAATCTAGTGGAGAACAACATACCCTCGCTCATCAATGG...CCAGATGTC 415
115 nglutrrtpttrpcysrthrgrludargalaleuasprtpgilyproaspclY. 131
416 AAGGCCCAACAGCAGGTTGTGCATATGTTGGTAGTATGTAACATCTCTCATC 465
132 .. GlyProAspLeuIleValAlaSpaspglyGlyAspAlaThrLeuLeuile 147
466 TCCAAAGGGCTTC.....GAATTCCAACACGCGGTCGCTTCC 503
148 HisgluglyAlaLysAlaaglueglutrryrslysSerGlyAlaLeuP 164
504 AGAGCCCAACGAAAGCTGACCAACCTCGAATACCGCGTCTTGTGCATAC 553
164 oAspProIlaSerThrAspSnalagluPhedInlleValLeuSerIle 181
554 TCAGCAGGCTCTCAACCAAGCAACAAGAACCACTGGCACACAGTTGCTGCC 603
181 leArgspsglyLeuLysSerAspPrometLysTyHHisLysMetLysasp 197
604 GGCATAACACGGTCTTCCGAGAGACAACAACAGGTGTCCACCGCCCTTA 653
198 ArgLeuValGlYalSerSerglugluthrThrThrGlyValLysArgLeuTy 214
654 CCAGCTTCGAGAGAGAGGGCAAACCTCTCTTCCCACCCATCAACGTCAACG 703
214 rglmetEglnghasnslglyThrLeuLeuPheProAlaIleAsnValAsna 231
704 ACAGCTTTACAAGATCCAAAGTTCGATACATCTACGCGTCCGCGCCACTCC 753
231 spserValThrLysSerLysPheAspAnleuTyGLyAsnArgHisSer 247
754 CTTATGATGCTATCAACCGTCTTCGATGTCATGATCGCGCGCAAGAC 803
248 LeuProAspLyLeuMetArgAlaThrAspValMetIleAlaGlyAsVa 264
804 AGCTTCGTCATGGGTTACGCGGANTGTCGGCAAGGGCTCGCTCAATCCC 853
264 lAlaLeuIleAlaGlyTyrgLYAspValGlySLySGlySalAlaAlam 281
854 TCCGTGGCCAAAGCGCTCGCGCTTATCATCACAACTCGACCCCAATCTGC 903
281 etLysGlnAlaGlyAlaArgValIleValThlgIuIleAspProIleCys 297
904 GCCTTCACAGCTGGCAGATGGAAGGCTACACAGTCCCGCGCATCGAGAGCT 953
298 AlaLeuGlnAlaTrmEglnGlyLeuGlnValLeuProLeuLinspva 314
954 CGTCAGAGATGTCATATCTTCGTATACATGCACAGAAACTGCGATATCA 1003
314 lValSerGluValAspIlePheValThrThrThrGlyAnLysAspIleI 331
1004 TCTCTGTTCATATATGCCCCACAGATGAAGAGTAAGGCTATTGTCGGTAAC 1053
331 leMetValSerAspMetArgLysMetLysAsnAsnAlaIleValCysasn 347
1054 ATGGCGCATCTCGAATACGAATATGATACAGATGCGCTCATGAATACC 1103
348 IlleglylHisPheAspAnsluIleAspMetLeuGlyLeuGluThrTyPr 364
1104 AGGCATCAACGACATCCCAATTCACGCGCAAGATACGACATGTGGAAATTC 1153
364 oglyValLysArgIleThrIleLysProGlnThrAspArgTyPrValPheP 381
1154 CAGAT...GGCGAGCATATCCGCTCTTGGAGAGGGCGCGCTTCAAC 1200
381 romspmrhglyArgGlyIleIleIleIleValAlaGluGlyArgLeuMetAsh 397

[illegible]

```

seq_name: /cogn2_6/ptodata/1/Paa/US09_1MW_COMB.pep:US-09-791-537-103494
seq_documentation_block:
  : Sequence 103494 Application US/09791537
  : GENERAL INFORMATION:
  :   APPLICANT: Biomimix, Inc.
  :   APPLICANT: Debe, Derek
  :   TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES
  :   TITLE OF INVENTION: METHODS OF USE THEREOF
  :   FILE REFERENCE: 261/210
  :   CURRENT APPLICATION NUMBER: US/09/791,537
  :   CURRENT FILING DATE: 2001-02-22
  :   NUMBER OF SEQ ID NOS: 153055
  :   SOFTWARE: PatentIn version 3.0
  :   SEQ ID NO 103494
  :   LENGTH: 485
  :   TYPE: PRT
  :   ORGANISM: Arabidopsis thaliana
  : US-09-791-537-103494

```

```
alignment_scores:      Quality: 1537.00      Length: 487
                       Ratio: 3.901      Gaps: 4
Percent Similarity: 80.903      Percent Identity: 62.012
```

```
alignment_block:
US-09-759-990-1 x US-09-791-537-103494 .
```

Align seg 1/1 to: US-09-791-537-103494 from: 1 to: 485

[illegible]


```

69 CGGCCGTGAAGAACTTACCTGTGAGAGAAATGCCAGTCTTATGG 118
   :||||| :||||| :||||| :||||| :||||| :|||||
46 eGlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMet 63
   :||||| :||||| :||||| :||||| :||||| :|||||
119 TTCTGCGTGAAGCGTTATTCGCTTCAAGCATGAGAGGTGCAGATC 168
   :||||| :||||| :||||| :||||| :||||| :|||||
63 LacysArgThrGluPheGlyProSerGlnProPheGlyAlaArgIle 79
   :||||| :||||| :||||| :||||| :||||| :|||||
169 TCGGTCCTCCCACTGACATGACAGAGCGGCTCTCATCGAGACACT 218
   :||||| :||||| :||||| :||||| :||||| :|||||
80 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThr 96
   :||||| :||||| :||||| :||||| :||||| :|||||
219 CACAGCTCTTGCTGATGATGATGAGGCTTCCCTCAACATCTTCTTA 268
   :||||| :||||| :||||| :||||| :||||| :|||||
96 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 113
   :||||| :||||| :||||| :||||| :||||| :|||||
269 CACAGATACAGCGCGTGTGCTATGCTGTGCGCCCAACAGCACCA 318
   :||||| :||||| :||||| :||||| :||||| :|||||
113 hrgInaSPHisAlaAlaAlaIleAlaAlaArgAspSerAlaAla.... 127
   :||||| :||||| :||||| :||||| :||||| :|||||
319 GAGAACGACCGCGTATCCCACTCTTCGCTGAGAGCGGCAACACTCC 368
   :||||| :||||| :||||| :||||| :||||| :|||||
128 .....ValPheAlaTrpIlysgIlyuThrLeuG 137
   :||||| :||||| :||||| :||||| :||||| :|||||
369 AGAATAGTGGAGAACACATACCGGCTCTCACATGCGCAGATGTCAG 418
   :||||| :||||| :||||| :||||| :||||| :|||||
137 ngIuTyTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyIlyg 154
   :||||| :||||| :||||| :||||| :||||| :|||||
419 GCCCACAGCAGGTGTGCATGATGTTGGTGTGATGCTACACTCCATCTCC 468
   :||||| :||||| :||||| :||||| :||||| :|||||
154 LyrProAspLeuIleValAlaSPHisArgIlyAspAlaThrLeuLeuHis 170
   :||||| :||||| :||||| :||||| :||||| :|||||
469 AAGGCGTTCGAA.....TTCGAAACAGCGGCTGTGTTCCAGA 506
   :||||| :||||| :||||| :||||| :||||| :|||||
171 GluGlyAllyAlaGluGluIlePheGluTyThrGlyGlnValProAs 187
   :||||| :||||| :||||| :||||| :||||| :|||||
507 GCCACAGAAAGCTGACAACTCGAATACCGGTGCTTGTGCTACACTCA 556
   :||||| :||||| :||||| :||||| :||||| :|||||
187 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLe 204
   :||||| :||||| :||||| :||||| :||||| :|||||
557 AGCAGGCTTCAACCAAGACAAAGAACACTGGCACACAGTGTGCGCGC 606
   :||||| :||||| :||||| :||||| :||||| :|||||
204 ysgIuGlyLeuGlnValAspProLysLystrHisLysMetLysGluArg 220
   :||||| :||||| :||||| :||||| :||||| :|||||
607 ATGAACGCTGTTCGAGAGAGACAAACAGAGTGTCCACCGCTCTACCA 656
   :||||| :||||| :||||| :||||| :||||| :|||||
221 LeuValGlyValSerGluGluThrThrArgIlyValLysArgLeuTyrg 237
   :||||| :||||| :||||| :||||| :||||| :|||||
657 GCTCGAGAGAGGAGCAACTCTCTTCCAGCATCAACGTCAACAGCAG 706
   :||||| :||||| :||||| :||||| :||||| :|||||
237 nmetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 254
   :||||| :||||| :||||| :||||| :||||| :|||||
707 CTGTCAAGTCCAGTTCGATTAACATCTAGCGGTGCCGCCACTCCCTT 756
   :||||| :||||| :||||| :||||| :||||| :|||||
254 eValAlThrLysSerLysPheAspAsnLeuTyrgLysArgHisSerLeu 270
   :||||| :||||| :||||| :||||| :||||| :|||||
757 ATCGATGATCAACCGTGTTCGATGATGATGATGATGATGATGATGAT 806
   :||||| :||||| :||||| :||||| :||||| :|||||
271 ProAspLysLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 287
   :||||| :||||| :||||| :||||| :||||| :|||||
807 TCTCGATATGGTTACGGCGATGTGGCAGAGGCTGCGCTCAATCCCTCC 856
   :||||| :||||| :||||| :||||| :||||| :|||||
287 aValIleCysGlyTyrgIlyAspValGlyLysGlyCysAlaAlaIleMet 304
   :||||| :||||| :||||| :||||| :||||| :|||||
857 GTGGCCAAAGGCGTCCGTTATCATCACAGAACTGACCAATCTGGCT 906
   :||||| :||||| :||||| :||||| :||||| :|||||
304 ysthAlaGlyAlaArgValIleValIleGluIleLeuProIleCysAla 320
   :||||| :||||| :||||| :||||| :||||| :|||||
907 CTCACGAGCTGCATGAGAGCTACAGAGTCCCGCGCATGAGAGAGTGT 956
   :||||| :||||| :||||| :||||| :||||| :|||||
321 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGlnAspValVa 337

```

```

957 CAAGGATGTCATATCTTGTTCATGATGACAGGAATCGCATATCATCT 1006
   :||||| :||||| :||||| :||||| :||||| :|||||
337 LserGluAlaSPHisPheValThrThrThrGlyAsnLysAspIleLeu 354
   :||||| :||||| :||||| :||||| :||||| :|||||
1007 CTGTTCACATGATGCGCCAGATGAAGGATTAAGCTTATTCGTTACATC 1056
   :||||| :||||| :||||| :||||| :||||| :|||||
354 eValAlaSPHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 370
   :||||| :||||| :||||| :||||| :||||| :|||||
1057 GGCACCTTCGATTAACGAATGATGATGATGATGATGATGATGATGAT 1106
   :||||| :||||| :||||| :||||| :||||| :|||||
371 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrrProG 387
   :||||| :||||| :||||| :||||| :||||| :|||||
1107 CATCAAGCACAATCCCAATCAAGCAGAAATACGATGTGGAAATCCAG 1156
   :||||| :||||| :||||| :||||| :||||| :|||||
387 YValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 404
   :||||| :||||| :||||| :||||| :||||| :|||||
1157 ATGGCCACAGCT...ATCGCTCTTTCGTGAGAGGCGCGCTTCTTAACCTT 1203
   :||||| :||||| :||||| :||||| :||||| :|||||
404 LuThrLysAlaGlyIleIleValAlaLeuAlaGluIlyArgLeuMetAsnLeu 420
   :||||| :||||| :||||| :||||| :||||| :|||||
1204 GGCCTCCCTACAGGTCACCCATCTTTCGTTATGTCATGTCATTCACAAA 1253
   :||||| :||||| :||||| :||||| :||||| :|||||
421 GlyCysAlaThrGlnHisProSerPheValMetSerCysSerThrPheAs 437
   :||||| :||||| :||||| :||||| :||||| :|||||
1254 CCAGACACTGCTCAGCTCAGCTCAGCTCAGCAAAAGAA.....GGAAATC 1297
   :||||| :||||| :||||| :||||| :||||| :|||||
437 ngInValIleAlaGlnLeuGluLeuThrProLysGlnLysAlaSerGlyLysT 454
   :||||| :||||| :||||| :||||| :||||| :|||||
1298 TCGAGAGAGAGTTPACACACTTCGAAAGCATCTGCATGAGAGAGTGCCT 1347
   :||||| :||||| :||||| :||||| :||||| :|||||
454 yrgLysLysValTyrgValLeuProLysHisLeuAspGlyLysValAla 470
   :||||| :||||| :||||| :||||| :||||| :|||||
1348 CGCCTCCACCTCGGATCTCTGATGTCCACTTACAAAGCTTACACAGAA 1397
   :||||| :||||| :||||| :||||| :||||| :|||||
471 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 487
   :||||| :||||| :||||| :||||| :||||| :|||||
1398 GCAGGCTGACTACATCAACGTTCCAGTTGAGGCTCTTACAGATGTGATG 1447
   :||||| :||||| :||||| :||||| :||||| :|||||
487 pGln**AspTyrgValSerIleProIleGluGlyProTyrgLysProProH 504
   :||||| :||||| :||||| :||||| :||||| :|||||
1448 CTTCACGTTAT 1458
   :||||| :||||| :||||| :||||| :||||| :|||||
504 lstyArgTyrg 507
   :||||| :||||| :||||| :||||| :||||| :|||||

```

seq_name: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-49252

seq_documentation_block:

```

: Sequence 49252, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791.537
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: Patent version 3.0
: SEQ ID NO 49252
: LENGTH: 485
: TYPE: PRT
: ORGANISM: Nicotiana sylvestris
US-09-791-537-49252

```

alignment_scores:

```

Quality: 1533.00 Length: 482
Ratio: 3.951 Gaps: 4
Percent Similarity: 80.498 Percent Identity: 63.278

```

alignment_block:

US-09-759-990-1 x US-09-791-537-49252

Align seg 1/1 to: US-09-791-537-49252 from: 1 to: 485

34	GACTAGCAAGATTGGCGACATCAACCTGCAGATGTTGGCGCGCTTAAGCACT	83
13	GLTYTLYSVALLYSAPRMESELGLNLAASPRHEGLYAGLYLEUGLITL	29
84	TACCCCTGTCGAGAAAGGAAATGCCAGGCTTATAGTTGTTCTGAGCGCT	133
29	EGULIEVLAAGLYVALGLIMETPROGLYLEUMETALACYSATRGTHRGLP	46
134	ATTCCGCTTCTAACCCATTGAAGGTGTCAAATCTCGAATCTCGTCCCTCAC	183
46	HEGLYPROSEGLINPRPHELYSGLYALATLSLTLEHGLYSELEUNIS	62
184	ATGACAGCTCCAGACAGCGCGTCTCATCGAGACACATCCAGCTCTTGATGC	233
63	MEETHRLIEGLNTHRLAVALLEUILEGLNTHRLIEUHLRLALEUGLYAL	79
234	TGATGTGATGAGGCTTCCTCCAACTGTTCTCTACAAATACAGCGC	283
79	AGLYVALATRGTPYSSERTCYSASTLEPHESETHRLINSPHLSIALA	96
284	CTGCTGCTATGTTGTCGGCCCAACAGGCGACACACAGAAAGCCACCGCT	333
96	LAHALAALIEALALARGASPERIALAIA.....	105
334	ATCCCACTCTTCGCCCTGGAGAGGCGCAACATCCCAATACTGGGACAA	383
106VALPHEALATRGLYGLYTHRLIEUGINGLYTTRTPCY	120
384	CACATACCGGCTCTCATATGGCCAGATGGTCAAGGCCACAGCGAGTTG	433
120	STHGLIATRGALALEUASPTRGILYPROGLYGLYGLYPROASPLEUILEV	137
434	TCGATGATGTGGTGATGCTACATCTCTCATCTCCAAAGGCTTC.....	477
137	ALASPRASGLYGLYASPRALATHRLIEUILLIEGLIUYALLYVALA	153
478GAATTCGAAAGACCGCGCTCGTTCACAGGCCAAAGCAACTGA	521
154	GLUGLUGLUPHEALALYALASNGLYTHRLIEPROASPRONASERTHRAS	170
522	CAACCTCGAATACGCGTCTGTTGCTACACTCAAGCAGGCTTCAAC	571
170	PASNALAGLUPHEGLINLEUVALLEUHLRLIEUYSGLUSELEUYS	187
572	AAGACAAACACACATGGGACACAAAGTTCGCGCGGATGAAGGCTGTTCC	621
187	HRASPRLEUULYSTYTHRLYSMELYSGLYARGLEUVALGLYVALSER	203
622	GAAGACACAAACACAGTCTCCACGCGCTTACACAGCTCGAGAAAGAGG	671
204	GLUGLUGLUPHRLTHRGLYALYALARGLEUUGLIMESGLINLAASNGI	220
672	CAAACTCTCTTCCACGCAATCAAGCTCAACAGCGCTGTATCAAAATGCA	721
220	YTHRLIEUHLPREROLALIASNVALASNSPSEVALTHRLYSELT	237
722	AGTTGATGATCACTTACAGGCTCGCGCACTCTTATCGATGATATCAC	771
237	YSPHEASPRASMLEUUGLYCYARGNISSELEUFGROASRGLYLEUMET	253
772	CGTGTCTCCGATGTCAATCGCGGCGCAACAGCTCTGTCAAGGGTTA	821
254	ARGALATINASRYVALMETLEIALAGLYVALALEUVALALAGLYT	270
822	CGGCAATGTCGCGCAAGGCTTGCGCTCAATCCCTCGTGGCGCAAGCGCTC	871
270	RELAYSAPVALGLYLYSGLYCYSAALALALALEUYSGLINALAGLYALAA	287
872	GGGTATATATCACAGAACTCGACCAATCTGGCGCTCTCCAGGCTGCGATG	922

```

287  rgvalallevalthrglulileasprolilecylalaleuqinalaithrmet 303
922  GAAGCGTACCAAGTCCGCCATCGAAGAAGTGTCAAGATGTCATAT 971
304  GluGlyLeuGlnValLeuThrLeuGlnAspValValSerAspValAspIle 320
972  CTGGGTACATGCACAGGAACCTGCATATCAATCTCTGTTGACATGATGG 1021
320  ephevalthrthrthrcllyslnlysaaspillelmetvalasphismeta 337
1022  CCCAAGTAAAGATTAAGGCTATGTGTGGTAAATCCGACATCGCCACTTCGATAC 1071
337  rglYsMetIysAsnAsnAlaIleValcysAsnIleGlylnAspPheAspAsn 353
1072  GAATGTGATACAGATGGCCTCATGAATACCCGACGATGACAGACATCCC 1121
354  GluIleAspMetLeuGlyLeuGlnThrTyrTrpGlyValIlylsArgIleTh 370
1122  AATCAAGCCAGAATACGACATGTGGAAATGCCAGATGCCACGCT..A 1168
370  rIleYsProGlnThrAspArgTrpValPheProAspThrAsnSerGlyI 387
1169  TCTCTCTCTTCTGCTGAGGCGCGCTTCTTAACTTGGCTGCTGCTACAGCT 1218
387  leIleValIleuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219  CACCCATCTTGTGTTATGTCAATGTCAATTCATCACAACAGACACATCGCTCA 1268
404  hIspProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420
1269  GCTCGACCTCTACGAAAGAAGA.....GGAAATCTCGAGAGAAGGTTT 1312
420  nleuGlnIleuTrpAsnGluIlySerSerGlyLysTyrGlnIlylsValT 437
1313  ACACACTTCCGAAGCATCTCGATGAAGAAGTGCCTGCTCCACCTCGGA 1362
437  yValAlaLeuProLysIleuAspGluIlyValAlaAlaLeuIleuIleuGly 453
1363  TCTCTCGATGTCCACCTTACAAAGCTTACACAGAAGAGGCTGACTATAT 1412
454  LysLeuGlnIyAlaIlylsLeuThrLysLeuSerLysAspGlnAlaAspTyrI 470
1413  CAACGTTCAGTGAAGGTCCTTACAAAGTGTGATGCTTACCGTAT 1458
470  eSerAlaProValGluGlyProGlyLysProAlaIleHisTyrArgTyr 485
seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB pep:US-09-791-537-37086
seq_documentation_block:
: Sequence 37086, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791, 537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37086
: LENGTH: 485
: TYPE: PRT
: ORGANISM: Trilicium aestivum
: US-09-791-537-37086

```

```

seq_name: /cgn2.6/picodata/1/paa/US09_NEW_COMB pep:US-09-791-537-37086
seq_documentation block:
: Sequence 37086, Application us/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMIL
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37086
: LENGTH: 465
: TYPE: PRT
: ORGANISM: Triticum aestivum
US-09-791-537-37086

```

```
alignment_scores:
  quality: 1524.00      length: 487
  ratio: 3.868          gaps: 4
Percent Similarity: 80.903  Percent Identity: 61.807
```

alignment block:
US-09-759-990-1 x US-09-791-537-37086 ..

Align seg 1/1 to: US-09-791-537-37086 from: 1 to: 485

```

19 ACTGGTCTCCATTCAGTACAGAAATTCGCGAGTCAACCTCCATCTCTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTyrTyrValLysAspLeuPheGlnAlaAspPh 24
69 CGCGCGTAAGAACTTACCTTGTCTGAGAAAGAAATGCCAGTCTTATGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMetA 41
119 TTCTGTGAGAGGTTATTCGGCTTCTAAGCATTTGAAGGTTGTCAGAAATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 lAcysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
169 TCTGGTCCCTCCACATGACAGTCAGACAGCGCTGCTCATCGAGCAACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrIle 74
219 CACAGCTCTTGGTGTGATGATGATGATGGCTTCTGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTyrPcysSerCysAsnIlePheSers 91
269 CACAAGATACAGCGCGTGTGCTATCGCTGTCGCCGCCCAACAGCACCA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 eGlnHisPheHisAlaAlaAlaAlaAlaAlaArgAspSerAlaAla.... 105
319 GAGAAGCCAGCGCGTATCCAGTCTTCGCTGGAAGGCGCAACACTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIleTyrPcysGlyGluThrLeuGlu 115
369 AGATATCTCGAGAAACATATACCGCGCTCTACATGGCCAGATGTCAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 uGluTyrTrpPcysThrGluArgCysLeuAspTrpGlyValGlyGly 132
419 GCCCAGCAGAGTGTGATGATGGTGGTGGTGTACTACCTCCATCTCTCC 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 LyrProAspLeuIleValAspAspGlyGlyAlaThrLeuLeuHis 148
469 AAGGCGCTTC.....GAATTGCAAAACGCGGTGCTGTGTGCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 GluGlyValLysAlaGluGluGluPheGluLysSerGlyLysValProAs 165
507 GCCAAGCAAGACTGACATCCGATACCGCGCTCTTCTGTACTACTCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 pProGluSerThrAspAsnProGluPheLysIleValLeuThrIleIle 182
557 AGCAGTCTTCAACCAAGCAAGCAACACTGCGCACAGTGTGCTGCGCG 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 rGAspGlyLeuLysThrAspAlaSerLysTyrArgLysMetLysGluArg 198
607 ATGAACGGTGTTCGGAAGACAAACAAGGTGTCCACCGCTTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrGly 215
657 GCTCGAAGAGGAGGCAAACTCTCTTCCACGACATCAAGCTCAAGCAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 nMetGlnGluSerGlyThrLeuLeuPheProAlaIleAsnValAsnAspS 232
707 CTGTACAAAGTCCAAAGTTCATACATCTACGCGTGCAGCACCCTT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 eValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
757 ATCGATGGTATCAACGCTGCTTCGATGTCATGATCGCGCAAGACAG 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyValAla 265
807 TCTCGTCATGGGTTACGCGCATGCGCGCAAGGCTCGCTCATCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 aValValCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaAlaLeu 282

```

```

857 GTGGCCAAAGCGCTGCGCTTATCATCAGACAACTGACCAATTCGCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 yGlnAlaGlyAlaArgValIleValIleThrGluIleAspProIleCysAla 298
907 CTCGAGGCTGCCATGGAAGGCTTACCGAGTCCGCCGATGAGGAAGTCGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLeuGluAspValVa 315
957 CAAGATGTCGATCTCTGTTACATGCACGGAAGAACTGCATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lSerGlnAlaAspIlePheValIleThrThrGlyAsnLysAspIleIleM 332
1007 CTGTGACATGATGCGCCAGATGAAGATGAAGATTAAGCTATTCGGTAACATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 eValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCGCATCTGATATACGAATTTGATACAGATGGCTCATGAATACCCAGG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluIleThrTyrProGly 365
1107 CATCAAGCACATCCCAATCAAGCCAAATACAGATGCGGGAATCCAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrPvalPheProG 382
1157 ATGGCCAC...GCTATCTCTCTTCTGCTGAGGCGCGCTTCTTAACCTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382 lThrLysThrGlyIleIleValLeuAlaGluGlyAlaGlyLeuMetCAsnLeu 398
1204 GGCTGGCTACAGATGACCATCTTCTGTTATGTCATATGCATTCACAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACTGCTCAGCTGACCTCTACGAAAGAGA.....GGAATTC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 nGlnValIleAlaGlnLeuGluLeuTyrPasnGluLysAlaSerGlyLysT 432
1298 TCGAGAAGAAGGTTACACACTTCGAAAGCATCTCGATGGAAGATCGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 yGluLysLysValTyrValLeuProLysHisLeuAspGlyLysValAla 448
1348 CGGCTCCACCTCGGATCTCGATGTCACCTTACAAAGCTTACACAGA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysSe 465
1398 GCAGGCTGATCATCAATCAAGTTCAGTGGGCTCTTACAAAGCTGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 rGlnSerAspTyrIleSerIleProIleGluGlyProTyrLysLeuArgL 482
1448 CTTACCCTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 euTyrArgTyr 485

```

seq_name: /cgn2.6/plotdata/1/paa/US09_NEM_COMB.pep:US-09-791-537-23700

seq_documentation_block:

- Sequence 23700, Application US/09791537
- GENERAL INFORMATION:
- APPLICANT: Biomix, Inc.
- APPLICANT: Debe, Derek
- APPLICANT: Danzer, Joseph
- TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
- FILE REFERENCE: 261/210
- CURRENT APPLICATION NUMBER: US/09/791,537
- CURRENT FILING DATE: 2001-02-22
- NUMBER OF SEQ ID NOS: 153055
- SOFTWARE: Patent version 3.0
- SEQ ID NO 23700
- LENGTH: 485
- TYPE: PR
- ORGANISM: Catharanthus roseus

US-09-791-537-23700

alignment_scores:

Quality: 1520.00	Length: 488
Ratio: 3.848	Gaps: 6
Percent Similarity: 80.943	Percent Identity: 62.295

alignment_block:
US-09-759-990-1 x US-09-791-537-23700 ..

Align seg 1/1 to: US-09-791-537-23700 from: 1 to: 485

```

19  ACTGTGCTCCATTCGAGTACAGAAATGCCGACATCACTCATGTTCT 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  8  ThreSerSerGlyArgGluTyrGlyValLysAspMetSerGlnAlaAsp 24
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 69  CGGCCGTAAAGAACTTACCTTGTGAGAAAGAAATGCCAGGCTTATG 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 24  eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuMet 41
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119  TTCTTGTGAGCGTTATTCGCTTGTATAGCCATTGAAGGCTGTCAGAAC 168
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 41  eGlyArgAlaGluIleGluLeuAlaGluValGluMetProGlyLeuMet 57
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169  TCTGTGCTCCATTCGAGTACAGAAATGCCGACATCACTCATGTTCT 218
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 58  ThrGlySerLeuHisMetThrIleGlnThrAlaValIleuIleGluThr 74
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219  CAGAGCTGTGCTGATGTCATGAGGCTTCTGCAACATCTTCTCTA 268
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 74  uThrAlaLeuGlyAlaGluValArgTyrPcysSerGlyAsnIlePheSer 91
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269  CACAGATACAGCCGCTGCTGATGCTGCTGCGGCCCAAGGCAAGCACACA 318
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 91  hGlnGlnHisAlaAlaIleAlaIleAlaArgAspSerAlaIle..... 105
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319  GAGAACCCAGCCGGTATCCAGCTTGTGCGCTGCAAGGCGCAAAACATGCC 368
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106  .....ValPheAlaIleTyrPcysGlyGluThrLeuGlu 115
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369  AGAATATGCTGGAGAACATACATACGCGCTCTCATG...CCAGATGCTC 415
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115  nGluTyrTrpTyrPcysThrGluArgAlaLeuAspTrpPcysProAspGly 131
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416  AAGGCCACAGCAGGTTGTCGATGATGAGTGATGCTACACTCTCTATC 465
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132  ..GlyProAspLeuIleValAlaAspArgLysAlaThrLeuLeuIle 147
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466  TCCAGAGGCTTC.....GAATTCGAAACAGCCGGTCTGTCTCC 503
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148  HisGluGlyValIleLysAlaGluGluTyrLysLysAsnGlyAlaLeuPr 164
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504  AGAGCCAAACAGAGTACAACTCGAATACCGCTGCTCTTCTCTACAC 553
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164  oAspProSerSerThrAspAsnAlaGluPheGlnIleValIleuThrIle 181
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554  TCAACGAGGTCTTCAACCAAGACAGAACCCATGGCACAGATTGTCGCC 603
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181  LeuArgAspLysLeuLysSerAspProThrLysTyrThrArgMetLysGlu 197
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604  GGCATGAACGGTGTTCGAGAGACAAACAAGGCTCCACCGCTCTA 653
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198  ArgLeuValIleGlyValSerGluIleThrThrGlyValLysArgLeuTyr 214
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654  CCAGCTCAGAGAGAGGCAAACTCTCTTCCAGCCATCAACGTCANAG 703
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214  GlnMetGlnAlaAsnGlyThrLeuLeuPheProAlaIleAsnValAsnA 231
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704  AGCGTGTACAAAGTCAAGTTCGATTAATATACAGGCTGCCCTCCTCC 753
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231  sPseValThrLysSerLysPheAspAsnLeuTyrLysAlaGlnHisSer 247
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
754  CTATTCATGATGATCAACCGCTCTCCAGATTCATGATGCGCGCAAGAC 803
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

248  LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 264
804  AGCTTCGTCATGAGGTTACGGCGATGTCGGAAAGGCTGCGCTCAATCC 853
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264  AlaValAlaValAlaGlyTyrGlyAspValGlyLysGlyAlaAlaIle 281
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854  TCCGTGGCCAAAGCCGCTCCGCTTATCATCACAGAACTGACCAATCTGC 903
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281  euLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
904  GCTTCACAGCTCCATGTAAGGCTACAGGCTGCGCCCATCGAGAAAT 953
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298  AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuThrLeuGluAspVa 314
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954  CGTCAAGATGTGATATCTTGTATCATGACAGCAAGAACTGCAATATCA 1003
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314  ValSerGlnAlaAspIlePheValThrThrGlyAsnLysAspIleI 331
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1004  TCTCTGTGACATGATGCCCAAGATGAAGATGAAGCTATTTGTGTAAC 1053
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331  LeuMetValAspHisMetLysMetLysAsnAlaIleValCysAsn 347
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1054  ATCGGCACCTTGATACGAAATGATACAGATGCGCTCATGAATGCC 1103
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348  IleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrTr 364
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1104  AGGCATCAAGCATCCCAATCAAGCCAGAAATCGACATGTGGAAATTC 1153
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364  oGlyValLysAlaGlyIleThrIleLysProGlnThrAspArgTyrAlaPhe 381
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1154  CAGATGCCACGCT...ATCTCTCTTGTGCTGAGGCGCGCTTCTTAC 1200
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381  roAspThrAsnSerGlyIleIleValLeuAlaGluGlyAlaGluMetAsn 397
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201  CTGTGCTGCTGATACAGGTCATCCATCTTGTGATATGATGATTCAC 1250
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398  LeuGlyCysAlaThrGlyHisPheSerPheValMetSerCysSerPheTh 414
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1251  AAACGACAGCATGCTGACGCTGACGCTGACGCTGACGAAAGAG...GGA 1294
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414  AsnGlnValAlaIleAlaGlnLeuGlnLeuIleThrPcysGluValGlyTh 431
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295  ATCTCGAAGAGAGGTTTACACACTTCCGAAACATCTCGATGAAGAAGTC 1344
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431  ystYrGluLysLysValIleValIleuProLysHisLeuAspGluVal 447
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345  GCTGCTCCGCTCCGATCTCTGATGTCCACCTTACAAAGCTTACACA 1394
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448  AlaAlaLeuHisLeuGlyLysLeuValAlaLysLeuThrLysLeuThrLy 464
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395  GAAGCAGGCTGATACATCAAGCTTCCAGTGAAGGCTCTTCAAGTCTG 1444
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464  sAspGlnAlaAspTyrIleSerValProIleGluLysProIlyLysProA 481
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1445  ATGCTTACCGCTTAT 1458
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481  LaHisTyrArgTyr 485
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

seq_name: /cgn2-6/plodata/1/paa/US09_NEW_COMB pep: US-09-935-625-2154

seq_documentation_block:

```

; Sequence 2154, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2154
; LENGTH: 467
; TYPE: PRT

```

```

; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..467
; OTHER INFORMATION: Ceres Seq. ID no. 3439073
US-09-935-625-2154

```

```

alignment_scores:
    Quality: 1518.00      Length: 470
    Ratio: 3.963          Gaps: 4
    Percent Similarity: 81.489    Percent Identity: 63.404

```

```

alignment_block:
US-09-759-990-1 x US-09-935-625-2154

```

```

Align seg 1/1 to: US-09-935-625-2154 from: 1 to: 467

```

```

70 GCCCGTAAAGAACTTACCTTGTGAGAAAGAAATGCCAGGTCTTATGGT 119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7  GlyAlaGluLeuLeuLeuAlaGluValGluMetProGlyLeuMetAl 23
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 TCTTGTGAGCGTTATTCGCTTCTAAGCCATGGAAGGTGTCAAGATCT 169
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 acysaIgrIthrGluPheGlyProSerGlnProPheIySGlyAlaIgrIle 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
170 CTGGTTCCTCCACATGACAGTCCAGACAGCCGCTTCATCGACAGACTC 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 hGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLeu 56
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 AAGAGCTGGTGGTGTGATGTCAGATGGGCTCCGCAACATCTTCTTAC 269
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 ThrAlaLeuGlyAlaGluValAlaGlyTrpCysSerCysAsnIlePheSerTh 73
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 ACAAGATACAGCCGCTGCTGTATCGTGTGCGCCCAAGCCACACACAG 319
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 rGlnAspHisAlaAlaAlaAlaIleAlaIleAlaArgAspSerAlaAla 87
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 AGAAGCCAGCCGGTATCCAGTCTTGGCTGGAAGGGGCAAACTCTCCA 369
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 .....ValPheAlaTrpLysGlyGluThrLeuGln 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
370 GAATATGGGAGAAACATACACATACCGGCTCTACATGGCCAGATGGTCAAG 419
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GluTrpTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 CCCACAGCAGGTTGTCATGATGATGGTGGTGAAGCTTACACCTCCATCTCCA 469
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 yProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHisG 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
470 AGGCGCTCGAA.....TTGAAACAGCCGCGTGTCTTCAGAG 507
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 IuGlyValIySalagIuGluIlePheGluIySerThrGlyGlnValProAsp 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 CCAACAGAGGTCGACACCTGGAATACCGCTGCTTCTTGTCAACATCA 557
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLeu 164
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
558 GCAGGTCTTCAACCAAGACAGAACACACACACACACACACACACAC 607
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 sGluGlyLeuGlnValAspProLysIySerThrIyHisIyMetIySGlyAla 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
608 TGAACGGTGTTCGAGAGACAGACACACAGTGTCCACCGCTTACACAG 657
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 euValIyGlyValIySerGluGluThrThrGlyValIyAspIyLeuIyGln 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 CTCGAGAAAGAGGCGCAACTCTCTCCAGCATCAAGCTCAAGCAGCG 707
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
198 MetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 214
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
708 TGTTCACAAGTCCAGTTCGATACATCTTACAGGCTCCGCGCTCCCTTA 757
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 rValIThrIySerIyAspPheAspAsnLeuGlyGlyCysArgHisSerLeu 231

```

```

758 TCGATGCTATCAACCGTGTCCGATGTCATGATCGCGGCAAGACACT 807
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 roAspIyLeuMetArgAlaThrAspValMetIleAlaGlyValAla 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 CTGCTCATGAGTTTACGCGCATGTGCGCAGAGGCTGCGCTCATCCCTCCG 857
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 ValIleCysGlyTyGlyAspValIyGlyGlyCysAlaAlaIleMetIy 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
858 TGGCCAAAGCGCGGTTTATCATCAAGAACTGACCAATCTGCGCTC 907
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 sThrAlaGlyAlaArgValIleValIThrGluIleAspProIleCysAla 281
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
908 TCCAGGCTGCGATGAAGGCTTACAGGCTCCGCGCATGAGAGACTGCT 957
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
281 euGlnIleLeuMetGluGlyLeuGlnValIleuThrLeuGluAspValVal 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
958 AAGAGTGTGATATCTCTGCTTACATGACAGAACTCGATATCATCTC 1007
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 SerGluAlaAspIlePheValThrThrGlyAsnIyAspIleIleMet 314
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1008 TGTTCACATGATGCGCCAGATGAAGATGAAGATATGTCGCTTACATCG 1057
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
314 tValAspHisMetArgIyMetIyAsnAlaIleValIyCysAsnIleG 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1058 GCCACTTCGATTAACGAATTGATACAGATGCGCTTATGAATACCAAGC 1107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
331 LysIAspAspAsnGluIleAspMetLeuGlyLeuGluIyThrTyTrpGly 347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1108 ATCAAGCAATCCCAATCAAGCAGATATACAGATGAGGAATTCACAGA 1157
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 ValIyAspTrpGlyThrIleIySerProGlnThrAspArgTrpValPheProG 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1158 TGGCCAGCT..ATCTCTCTTCTGCTGAGGCGCGCTTCTTACCTTG 1204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 uThrIyAlaGlyIleIleValIleuAlaGluIyArgIleMetIyAsnLeu 381
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1205 GCTGCGCTACAGGTCACCATCTTCTGATATGATGTCATGATCAACAC 1254
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
381 LysAlaIyThrGlyHisProSerPheValMetCysSerSerPheThrAsn 397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1255 CAGACACTGCTCAAGCTCGACCTTACGAAAGAGA.....GGAAATCT 1298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
398 GlnValIleAlaGlnLeuGluIyLeuTrpAsnGluIyAlaSerIyIySTy 414
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 CGAGAAAGAGTTTACACACTTCCGAAAGCATCGATGTCATGAGTGGCTC 1348
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 rGluIyLysValIyValIleuProLysHisLeuAspGlyValAlaAl 431
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1349 GCGTCCACCTCGATCTCTGATGTGTCACCTTACAAAGCTTACACAGAG 1398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
431 euLeuHisLeuGluIyLysLeuGlyAlaArgLeuThrLysLeuSerIyAsp 447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1399 CAGGCTGACTACATCAAGCTTCCAGTTGAGGCTTACAAAGTGTATGC 1448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 Gln***AspTrpValSerIleProIleGluGlyProTrpLysProIle 464
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1449 TTACCGTTAT 1458
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
464 sTyTrpTyTr 467

```

```

seq_name: /cgn2_6/plodata/1/paa/US09_NEW_COMB.pep:US-09-791-537-13836

```

```

seq_documentation_block:
; Sequence 13836, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791, 537

```


; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 13836
 ; LENGTH: 485
 ; TYPE: PRT
 ; ORGANISM: Mesembryanthemum crystallinum
 ; US-09-791-537-13836

alignment_scores:

Quality: 1518.00 Length: 487
 Ratio: 3.824 Gaps: 4
 Percent Similarity: 81.520 Percent Identity: 61.602

alignment_block:

US-09-759-990-1 x US-09-791-537-13836 ..

Align seq 1/1 to: US-09-791-537-13836 from: 1 to: 485

```

19  ACTGGTGGTCCATTCAGTACAGAAATGGCGACATCACTCCATGTTCT 68
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8  ThrSerSerGluArgGluThrGlyValLysAspMetSerGlnAlaAsp 24
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCGGTGAAGAACTTACCTGCTGAGAAAGAAATGCCAGTCTATGG 118
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24  eGluArgLeuGluLeuGluLeuAlaGluValGluMetProGluLeuMet 41
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTTCTTGGAGCGGTATTCGGCTTGAAGCCATTGAAGGGTGCAGAAATC 168
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41  IacysArgThrGluPheGlyProSerGlnProPheGlyAlaLysIle 57
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGGTGGTCCATTCAGTACAGAAATGGCGACATCACTCCATGTTCT 218
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58  ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTTGGTGGTATGATGATGATGATGATGATGATGATGATGATGAT 268
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74  uThrAlaLeuGluAlaGluValAlaGlyThrCysSerCysAsnIlePheSer 91
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAAGATACAGCCGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 318
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91  hGlnAspHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 105
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAAGCCAGCCGCTATCCAGTCTTCGCTGGAAGGCCGGAACACTGCC 368
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106  .....ValPheAlaThrPheGlyGluThrLeuGlu 115
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGATACATGGGAGAACATACCGCGCTTCACATGGCCAGATGGTCAAG 418
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115  nGluThrTrpTrpCysThrGluThrAlaLeuAspTrpGluAlaGlyIleG 132
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 GCCCAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 468
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132  LyrProAspLeuIleValAlaAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 AAGGCGCTTCT.....GAATTCGAAACAGCCGCTGCTGCTGCTGCTGCT 506
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149  GluGluValLysAlaGluGluGluGluGluGluGluGluGluGluGluGlu 165
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 GCCAAGACAGAGTGCACACCTGCAATACCGCTGCTGCTGCTGCTGCTGCT 556
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165  pProThrSerThrAspAsnProGluPheGlnLeuValLeuGluLeuIleHis 182
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557 AGCAGGTCTTCAACCAAGAACAGAACACATGGCAGACAGTGTGCGCGGC 606
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182  rGAspSerLeuLysValAspProLysArgGlyHisLysMetLysThrArg 198
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607 ATGAACGGTGTTCGGAAGAGACACACACAGTGTGCTCACCGCTCTTACCA 656
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199  LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrGlu 215
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657  GCTGAGAAAGAGGAGCAACTCTCTTCCAGCATCACTCAAGCAG 706
  
```

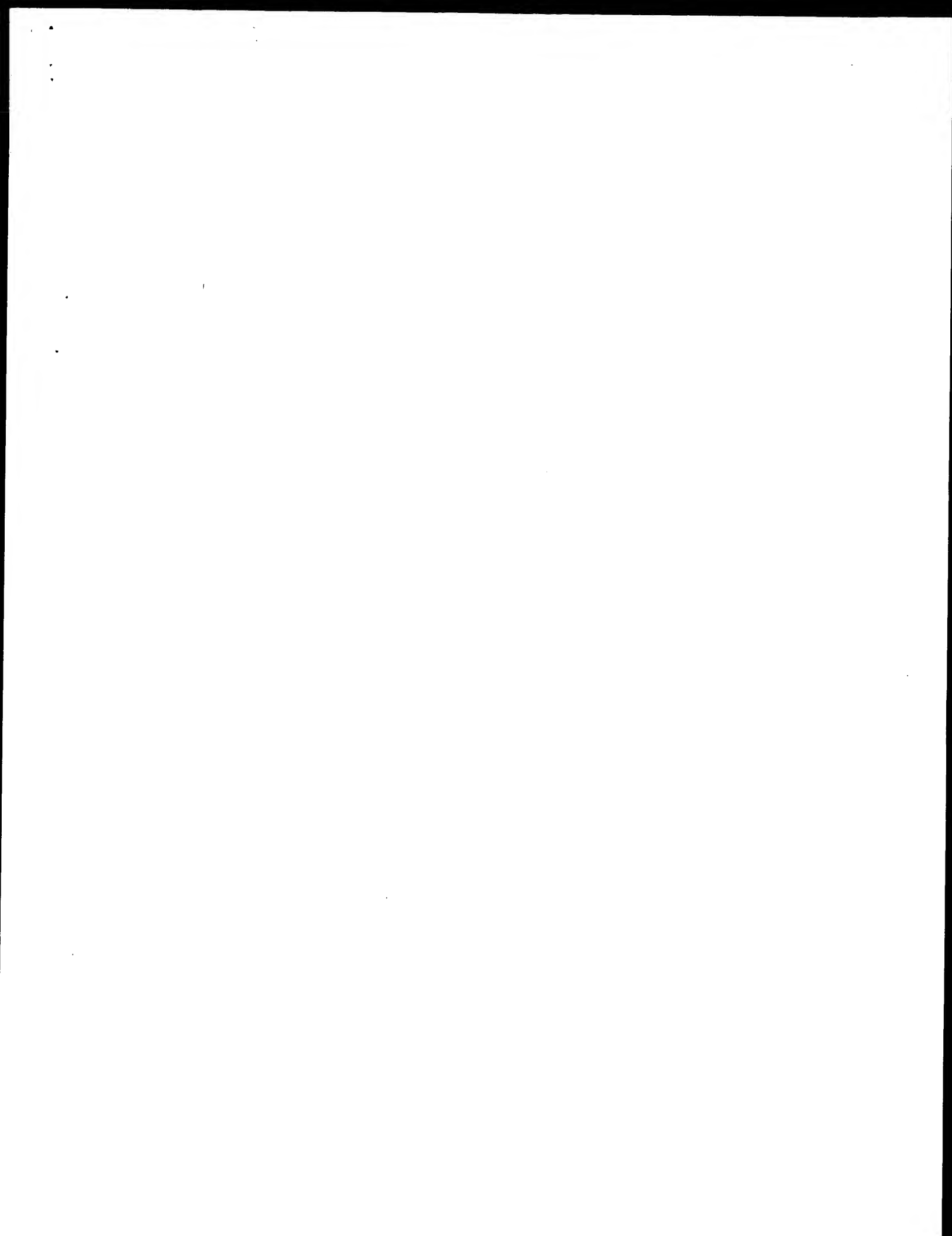
```

215  nMetGlnAlaThrGlyThrLeuLeuPheProAlaIleAsnValAsnAspS 232
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707  CTGTTCACAAAGTTCAGATACATCACTGACGCTGCGCCACTCCCTT 756
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232  erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757  ATCGATGGTATCAACCGCTCTCCATGTCATGATGATGATGATGATGATGAT 806
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249  ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValGlu 265
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807  TCTGTCATGGGTACGGGATGTCGGCAAGGGTGGGCTCAATCCCTCC 856
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265  yValValCysGlyTyrGlyAspValGlyLysGlyCysAlaLeuAlaLeuL 282
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
857  GTGGCAAGCGCGCTCGTTATCATGACAGAACCTGCAATCTGCGCT 906
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282  yAlaAlaGluAlaArgValIleValThrGluIleAspProIleCysAla 298
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907  CTCAGGCTGCGATGGAAGGTACACAGGTCCCGCCGATCGAAGATGCT 956
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299  LeuGlnAlaLeuMetGluGluPheGlnIleLeuThrLeuGluAspValIva 315
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957  CAAGGATGTCATATCTGTTGATTCATGCAAGAACTGGCATATCATCT 1006
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315  lSerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleIleM 332
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332  etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057  GGCCACTTGCATACAGAAATGTATACAGATGGCTCATGAATATCCAGG 1106
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349  GlyHisPheAspAsnGluIleAspMetLeuGluGluAsnTyrProGlu 365
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107  CATCAAGCAGCATCCCAATGAGCCAGAAATGCAATGATGAGAAATCCAG 1156
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365  yValLysArgIleThrIleLysProGlnThrAspArgPheValPheProG 382
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157  ATGGCCAGC...GCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382  LuThrAsnThrGlyIleIleValLeuAlaGluGluAlaGlyLeuMetLysLeu 398
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204  GCTCGCGCTACAGGTACCCATCTTCTGTTATGTCATGTCATTCACAAA 1253
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399  GlyCysAlaThrGlyHisPheSerPheValMetSerCysSerPheThrAs 415
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254  CGAGACACTGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1297
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415  nGlnValIleAlaGlnLeuGluLeuIleuTrpAsnGluThrAlaSerGlyLysT 432
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298  TCGAAGAAAGGTTTACACACTTCGGAAGCATTCGATGGAAGATCGCT 1347
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432  yGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348  CGCTCTCACCTGGATCTCGATGTCACCTTACAAAGTTTACACAGAA 1397
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449  AlalaLeuHisLeuGluLysLeuGluAlaLysLeuThrLysLeuSerLysAs 465
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398  GCAGGTGACTACATCAAGTTCAGTGAAGGCTTACAAAGTGTGATG 1447
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465  pGlnAlaAspTyrIleSerValProValGluGlyProTyrLysProAlaH 482
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448  CTTCACCGTTAT 1458
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482  lSTyrArgTyr 485
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: /cgn2_6/ptodata/1/paa/US09_NEW_COMB pep: US-09-935-625-738
seq_documentation_block:
; Sequence 738, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
  
```


Mon Aug 19 15:16:20 2002

us-09-759-990-1.n2p.rapn

Page 17



Db 301 ggggtgcgggtgttcggtggaagggcgaaacgtccgaagatactgtgtgtgcagga 360
 Qy 391 cgcgcctacatggtgcagatggtlcaagggccacacagaggtgttcgaatgtgtat 450
 Db 361 cagggcgtcactcgtgcggagcgccgaagggccacatgataccctcgaagagggcgga 420
 Qy 451 gctacacctcctacatcccaaggggttcgaatcgaacacgggtgtgttccagaca 510
 Db 421 gccaacatgtgttcccaaggggtgtcagtaagaaggggggtgtcccgaccgg 480
 Qy 511 acaagaagtcgaacactcgaatccggtgtgttctgttaacatcgaacaggttcaac 570
 Db 481 gataccgcgaagcagcaggtgttcgcatcgttctgtgaactgtctccgaagaccttgcg 540
 Qy 571 caagacaagaacacacgtgcacacagttgtctgcggcagatgaaaggtgttccgaagaca 630
 Db 541 gaaacaccccgcaagtggaagagatcgtctccggatcgaaggggtgtacccgaagagac 600
 Qy 631 acaacagtggtccacacggtctaccagctcgaagaagggcgaaactcctcccgacc 690
 Db 601 acgacgggggtgacccgctctacagatgcagcgttccggacagctgttcccgcg 660
 Qy 691 atcaaaagtcgaacgctgttacaagttccaagtgcataacatcagctcagctcgccac 750
 Db 661 atcaacgtlcaacgactccgctcagaaagagaagttcgaacaaagttacggtcgccgac 720
 Qy 751 tcccttatcgtatgtaacacgtgtctccgattcgtatgtaacgagcgaagacagctc 810
 Db 721 tgcgtatgcagcagcgtacaaacggcgccatgcagctgtctcagcgggaaggtcggtg 780
 Qy 811 gtaatggttgaagcgtatgcgaaggggtgtcgtcgaatccctcgtgtgccaagcgct 870
 Db 781 gtcgcccgtatcagcagcgtcgttgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtc 840
 Qy 871 cgcgttatcctacagaacactcgcacccacatcgtcgtcgtcgtcgtcgtcgtcgtcgtc 930
 Db 841 cgcgttatcctacagaacactcgcacccacatcgtcgtcgtcgtcgtcgtcgtcgtcgtc 900
 Qy 931 caggtcgcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 990
 Db 901 caggtcgcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
 Qy 991 aactcgtatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1050
 Db 961 aactcgtatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1020
 Qy 1051 aactcgtatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1110
 Db 1021 aactcgtatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1080
 Qy 1111 aagcacaaccac 1170
 Db 1081 aagcacaaccac 1140
 Qy 1171 cctccttctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1230
 Db 1141 atcgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1200
 Qy 1231 gttatgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1290
 Db 1201 gttatgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1260
 Qy 1291 ggaacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1350
 Db 1261 gaccagttacccgacggcggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1320
 Qy 1351 ctcacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1410
 Db 1321 ctcacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1380
 Qy 1411 atcaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1457
 Db 1381 atcgtgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1427

RESULT 2
 US-09-540-209b-2680
 ; Sequence 2680, Application US/09540209B
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOID
 ; FILE REFERENCE: 2709, 1001-001
 ; CURRENT APPLICATION NUMBER: US/09/540, 209B
 ; NUMBER OF SEQ ID NOS: 10444
 ; SEQ ID NO 2680
 ; LENGTH: 1500
 ; TYPE: DNA
 ; ORGANISM: B. fragilis
 US-09-540-209b-2680
 Query Match 37.4%; Score 547; DB 5; Length 1500;
 Best Local Similarity 63.4%; Pred. No. 2.5e-151;
 Matches 904; Conservative 0; Mismatches 485; Indels 36; Gaps 3;
 Db 37 tacagaattgcgacatcaacatccatcgttctcgcgcgtaaggaactacactgtcgtg 96
 Qy 112 tacaagtgagacatcaacgttgcgacttcggtcgcgaagaaatcgaactgtgcagaa 171
 Db 97 aagaatgtccaggtcctatgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 156
 Qy 172 aagaatgtccaggtcctatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 221
 Db 157 gtttcgaatcctgttctcctccatcgaatcgaacgagcagcgcgtccatcagaca 216
 Qy 232 gttgcgcgtatcagtgatcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 291
 Db 217 ctcaagctctgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 276
 Qy 292 ctgtgaacccctggtgagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 351
 Db 277 acaacgctgtgtatgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 336
 Qy 352 catgtcgtcagcgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 384
 Db 337 ccaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 396
 Qy 385 gccgtatttgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 444
 Db 397 ctcaatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 456
 Qy 445 ctgaacttgaagggggaagacagcagcagcagcagcagcagcagcagcagcagcagcagc 504
 Db 457 ctccatcctcgaaggggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 516
 Qy 505 atgatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 564
 Db 517 gctgacaactcgaatccgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 576
 Qy 565 catgcccgaagacgaatagaat---tgaatgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 621
 Db 577 aagaacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 636
 Qy 622 aagaacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 691
 Db 637 gttgtcagccgctcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 696
 Qy 682 ggtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 741
 Db 697 gtaacagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 756
 Qy 742 gtaacagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 801
 Db 757 atcgtgtatcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 816

Query Match	36.1%	Score 526.8;	DB 8;	Length 1407;
Best Local Similarity	62.5%;	Pred. No. 2.4e-145;		
Matches 897;	Conservative	0;	Mismatches 497;	Indels 42;
				Gaps 3

[illegible]

;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ. ID NOS: 47374
;; SEQ. ID NO 32437
;; LENGTH: 1398
;; TYPE: DNA
;; ORGANISM: Ralstonia metallidurans
US-60-360-039-32437

Query Match 32.8%; Score 481.2; DB 8; Length 1398;
Best Local Similarity 61.1%; Pred. No. 7.4e-132;
Matches 874; Conservative 0; Mismatches 508; Indels 48; Gaps 4;

```
OY 34 gagtacaagaattgcgcacatcaaccctcaatgcttctcgccgtaaggaattacccttgc 93
DB 10 gactacatctggtccgcatctgcgtccgctggtggtccggaagaattcgcacatgcgc 69
OY 94 gagaagaagaatgcgaagctctatgcttctctgtagagcttattccgcttctaagcattg 153
DB 70 gagacgcaaatgcccgtctgtagtgcattccgcaagaattccgcgcgcgcgcgcgcgcgc 129
OY 154 aagggtgtcagaattctctgttccctccacaatgacagtcagacagccgctctacag 213
DB 130 aagggtgtcagcgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 189
OY 214 aacatcacagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 273
DB 190 agcgtcaaggaagcagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtg 249
OY 274 gatacagcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 333
DB 250 gatcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 282
OY 334 atcccaagcttctgctcgtggaagggcgaagaacatcccaagaattctggagaaacatacgc 393
DB 283 agcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 342
OY 394 gctctacatgtgcagatgtctcaagggccacagcaggtgtgtgtgtgtgtgtgtgtgtgt 453
DB 343 atcttcgactgtggtgcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 402
OY 454 aacatccctcatctccaagggtcttgaattgaaacagcgcgtgtgtgtgtgtgtgtgtgtgt 513
DB 403 acgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 456
OY 514 gaaggtgacaacctcgatcaaccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 573
DB 457 aagcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 516
OY 574 gacaagaaccactgtgcacacagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 633
DB 517 gattcgacctgtgtacagcgcgaacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 576
OY 634 acaggtgtccacccgcctctcaacagctcgaggaagggcaaacctctcttcccaagcattc 693
DB 577 acggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 636
OY 694 aacgtcaacgacgctgttacaaggtcgaaggttcgaataacatcgaaggttcgaaggttc 753
DB 637 aacgtcaacgacgctgttacaaggttcgaaggttcgaataacatcgaaggttcgaaggttc 696
OY 754 ctcttcgattgtatcaaacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 813
DB 697 ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 756
OY 814 atgggtttagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 873
DB 757 ggcgtgttagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 816
OY 874 gtatcatcacaagaactctgacccaattctgctgtctcagcgtctgacgtggaagcttacc 933
DB 817 gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 876
```

```
OY 934 gtccgcgcacatcaggaaggtctcaaggtatgtcatctcttctgttacaatgcaggaac 993
DB 877 gtgtgtgacgaatgtgactacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 936
OY 994 tgcgataatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1053
DB 937 taccacgtgtatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 996
OY 1054 atgcgcacatctgtatcaagaattgtatcaagatgtgcctcactgaataaccagatcaag 1113
DB 997 atgtgcacatctgtatcaagaattgtatcaagatgtgcctcactgaataaccagatcaag 1056
OY 1114 cacatcccaatcaagaacgaatatacagcagcagcagcagcagcagcagcagcagcagcag 1173
DB 1057 atcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1107
OY 1174 ctcttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1233
DB 1108 atctgtgcaggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1167
OY 1234 atgtcaatgtatctacaacaacagacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1293
DB 1168 atgtcaatgtatctacaacaacagacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1227
OY 1294 aatctgcaggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1347
DB 1228 agcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1287
OY 1348 cgcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1407
DB 1288 cgcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1347
OY 1408 tacatcaacgttcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1457
DB 1348 tacatcaacgttcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1397
```

RESULT 13
US-60-360-039-28470
Sequence 28470, Application US/60360039
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianteng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360, 039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 28470
LENGTH: 1383
TYPE: DNA
ORGANISM: Burkholderia fungorum
US-60-360-039-28470

Query Match 32.8%; Score 479; DB 8; Length 1383;
Best Local Similarity 60.6%; Pred. No. 3.3e-131;
Matches 863; Conservative 0; Mismatches 520; Indels 42; Gaps 3;

```
OY 34 gagtacaagaattgcgcacatcaaccctcaatgcttctcgccgtaaggaattacccttgc 93
DB 1 gattcatgtgtgcgacatgtcgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 60
OY 94 gagaagaagaatgcgaagctctatgcttctctgtagagcttattccgcttctaagcattg 153
DB 61 gagaagaagaatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 120
OY 154 aagggtgtcagaatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 213
```


Db	448	aagccgacccaagaggaagaagtcgcgtctgtacaaagtcgatcagacatccatctgtgacgc	507
OY	574	gacaaagaaccaactcgtgcacacagttgtcgtccgcgcaltgaacaggtgtttccgaagaagaca	633
Db	508	gaccgcagcgtgtactctccacgcgccttgcgcacatcaagggcgtgtaccgaagaacaacag	567
OY	634	aaagtggtccacgcgcctctccacgcgtctgagagaagaggcaaatcctcttccacgcac	693
Db	568	accggcggttccacgcgtctgtactagatgtgaaaaaagagccgcgtgcgtttcccttgatc	627
OY	694	aacgttcaacgacgcgtgttacaagagttcccaagtttcgtatacactcaagcgtccgcgcac	753
Db	628	aacgttcaacgacccgtgttacaacaggttcgaagtctgacaacacccgttatggtccgtgtacg	687
OY	754	cttatctgatgtatcaacacgcgtgccttccgatgtcaatctgcgcgcgcgaagacgtctc	813
Db	688	ctgtctgcagcgcgtacacagc	747
OY	814	atgggtgtatccgcgcgtatctcgcgaaggggtgcgtccgtcaatccctccgcgtgcgaagcgtcgc	873
Db	748	gccgcgttaccgcgcgcgtgtggaagggcgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	807
OY	874	gttatcatcaacgaactcgcgcaccaatctgcgcctctccacgcgtcccatgtgaaggtaccag	933
Db	808	ggtgtggtgcacccaagaatcgcattccgatctgcgtatctgcgcgcgcgcgcgcgcgcgcgcgc	867
OY	934	gtccgc	993
Db	868	gtccgtgacgatatgtgaatacgc	927
OY	994	tgcgcgtatcatcctcgttgcacatgatctgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	105
Db	928	taccacgtgatccgc	987
OY	1054	atccgcgcacattcgataacgaagaattgatatacagaatggtccctcatgatataccacagcatcaag	111
Db	988	atccgtgtacattctgcacctcgcgaatacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	104
OY	1114	caacatcccaatccgaacgcagagataatcgaaatgtgggaattcccgatgtgcgcgcgcgcgcgc	117
Db	1047	-----catcaagccgc	109
OY	1174	cttctctgtgagggcgcgcgcctctttaaactcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	123
Db	1099	ctcgtcgtcgtgaagccgc	115
OY	1234	atgtcaattgtcatctacaacacgc	129
Db	1159	atgtcgtacacactcgtctacgcgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	121
OY	1294	aactctgagaaagaggttttacaacactcgcgaagcatctcgatgaagaaggtcgcgcgcgc	135
Db	1219	aaggtacgaagaacaggtgtgcgcgtcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	127
OY	1354	caacctcgc	141
Db	1279	catctcgtcgc	133
OY	1414	aagcgttccaggttgaggtgccttcaaatctcgtatcgtatccgttat	1458
Db	1339	ggcgcgttcgacagaagacgcgcgcgttcaagccgcgaacacactacgcgcgtat	1383

RESULT 15
US-60-360-039-40463
Sequence 40463, Application US/60360039
GENERAL INFORMATION:
APPLICANT: CAO, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.

```

: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
:
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
:
: FILE REFERENCE: 38-10(52052)A
:
: CURRENT APPLICATION NUMBER: US/60/360,039
:
: CURRENT FILING DATE: 2002-02-21
:
: NUMBER OF SEQ. ID NOS: 47374
:
: SEQ. ID NO 40463
:
: LENGTH: 1392
:
: TYPE: DNA
:
: ORGANISM: Caulobacter crescentus
:
: US-60-360-039-40463

```

Query Match	32.6%	Score 475.6;	DB 8;	Length 1392;
Best Local Similarity	60.4%	Pred. No. 3.3e-130;		
Matches 862;	Conservative 0;	Mismatches 524;	Indels 42;	Gaps 3;

[illegible]

Result No.	Score	Query Match	Length	DB	ID	Description
1	1461	100.0	1461	29	US-09-759-990-1	Sequence 11, Appl
2	586.4	40.8	1557	23	US-09-606-740A-219	Sequence 219, Appl
3	596.4	40.8	1557	29	US-09-746-660A-97	Sequence 97, Appl
4	594	40.7	1422	29	US-09-738-626-836	Sequence 836, Appl
5	536	36.7	1902	25	US-09-654-617-386668	Sequence 386668, Appl
6	536	36.7	1902	27	US-09-684-016-386668	Sequence 386668, Appl
7	535.4	36.6	1902	25	US-09-615-884-23	Sequence 23, Appl
8	531.6	36.6	1488	18	US-09-418-884-23	Sequence 7943, Appl
9	531.6	36.4	1434	58	US-09-703-708-7943	Sequence 7943, Appl
10	531.6	36.4	1434	55	US-60-164-320-7943	Sequence 7943, Appl
11	531.6	36.4	19798	27	US-60-183-791-7943	Sequence 691, Appl
12	531.6	36.4	19798	58	US-09-703-708-7943	Sequence 691, Appl
13	531.6	36.4	19798	55	US-60-164-320-691	Sequence 691, Appl
14	529.6	36.2	1396	57	US-60-183-791-691	Sequence 691, Appl
15	529.6	36.2	1396	23	US-09-606-740A-223	Sequence 223, Appl
16	512	35.0	1815	70	US-09-746-660A-101	Sequence 101, Appl
17	512	35.0	1937	25	US-09-654-617-260094	Sequence 2389, Appl
18	512	35.0	1937	25	US-09-654-617-260094	Sequence 260094, Appl
19	508.8	34.8	1865	28	US-09-708-427-77518	Sequence 260094, Appl
20	506.6	34.7	1458	78	US-09-708-427-22303	Sequence 77518, Appl
21	506.6	34.7	1458	75	US-60-361-294-15	Sequence 22303, Appl
22	506.6	34.7	1969	27	US-09-654-617-118096	Sequence 118096, Appl
23	506.6	34.7	1969	25	US-09-684-016-118096	Sequence 118096, Appl
24	506	34.6	1401	19	US-09-514-000-5970	Sequence 5970, Appl
25	506	34.6	1401	29	US-09-514-000-5970	Sequence 5970, Appl
26	506	34.6	1401	31	US-09-739-449-5970	Sequence 5970, Appl
27	506	34.6	25572	19	US-09-514-000-346	Sequence 346, Appl
28	506	34.6	374732	29	US-09-739-449-216	Sequence 216, Appl
29	506	34.6	374732	31	US-09-803-110-216	Sequence 216, Appl
30	497.8	34.1	1728	24	US-09-620-948-3581	Sequence 3581, Appl
31	496.8	34.0	1728	22	US-09-665-309A-47857	Sequence 47857, Appl

Db 1441 tctgactctaccgtlataa 1461

RESULT 2
US-09-606-740A-219
Sequence 219, Application US/09606740A
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
FILE REFERENCE: BGI-121CP
CURRENT APPLICATION NUMBER: US/09/606,740A
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142,101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148,613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187,970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931415.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931418.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931434.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931435.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931443.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931453.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931457.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931465.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931478.0
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931510.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931541.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931573.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931592.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931632.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931634.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932130.2

PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932186.8
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932206.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932227.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932228.7
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932229.5
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932230.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932926.5
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932928.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933004.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19940764.9
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940765.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940766.5
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940832.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941380.0
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941394.0
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941396.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942076.9
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942077.7
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942079.3
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942086.6
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942087.4
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942095.5
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942124.2
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942129.3
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 1158
SEQ ID NO 219
LENGTH: 1557
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1534)
OTHER INFORMATION: RXN00132
US-09-606-740A-219

Best Local Similarity 64.5%; Pred. No. 4.6e-155;
Matches 923; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 32 tcgagtagaattgcgcacatacaactcattcttcggtcgtgaaggtaactaccctg 91
DB 114 tggactcaaggttcgcgactcttaccatagcagggcagcgcgtccacagatctgtc 173
QY 92 ctgagaagaataatgcaggtctatagttcttcgtgagcggttattccgtcttaagcat 151
DB 174 cagaatattagatgcaggtctcatcagttctgcaaggaattccgacagacgacct 233
QY 152 tgaagggttcgaatctctgttccctccacatgacagttccacacagccgtccacacg 211
DB 234 tgaagggttcgacgtatctgttcttccatccacatgacaggtccacacgctgtctatg 293
QY 212 agacactcaagctctgttctgtatgttcagatgaggtctccgtcaacatctcttaac 271
DB 294 agacctcaactgttctgtgaggtgaggtctgtgaggtctccgtcaacatctcttaac 353
QY 272 aagatcacgagcgtctgtctatctgtctgtccacacagacacagagaagccagc 331
DB 334 aggatgaggtctgcaggtctgtctgtctgtccacacagacacagagaagccagc 410
QY 332 gtatcccaagctctgtcgtggaaggcgaacactccacagatattggaagaacatacc 391
DB 411 gtgtccagatctcgtgtggaagggtgaggtcaggtgaggtgaggtgaggtgaggt 470
QY 392 ggcgtctcaatggtccagatgtgtcaaggccacagcaggtgtgtcagatgtgtgtatg 451
DB 471 agatctcagctgtggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggt 527
QY 452 ctacactctccatccacaggtcttcgaatttcgaagaacagcgtgtcttcacagacaa 511
DB 528 tcaacacaggtctgtcttcgtggtcgtggtcgtggtcgtggtcgtggtcgtggtcgt 587
QY 512 cagaagctgcagacactcgaataccgctgtctgtctgtacacacaggtcttcacac 571
DB 588 aggcacacagatctcgaatgaggtatcgtcattctgtggtcgtggtcgtggtcgtggt 647
QY 572 aagacaagaacacactgcgcacacaggtgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 631
DB 648 cagagctgcgcagatgtggtggtcaggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 707
QY 632 caacaggtgtccacagcctctcaacagctcgaagaaggcgaacactctctccacagca 691
DB 708 ccacccgtgtgcacagcctctcaacagctcgaagaaggcgtcgtcgtcgtcgtcgtcgt 767
QY 692 tcaacgtcagcagcgtcttcaacaagttccagatccatcaacatcaggtcgtccact 751
DB 768 tgaacgtcagcagcgtcttcaacaagttccagatccatcaacatcaggtcgtccact 827
QY 752 ccctatcgtatggtatcaacaggtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 811
DB 828 ccctatcgtatggtatcaacaggtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 887
QY 812 tcatggttcaagcgtatcgcgaagggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 871
DB 888 tctcgtgttcaagcgtatcgcgaagggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 947
QY 872 ggcgttcatcatcaagaactcgaacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 931
DB 948 ggcgttcatcatcaagaactcgaacacatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1007
QY 932 aggtcgc 991
DB 1008 ctgtgtgtcgc 1067
QY 992 actgtgacatcatctgttgcacatgaggtccacagatgaggaaggaaggtatgtcgtgta 1051
DB 1068 acaagagcatcatcttctcgtcgaagatgaggtcgaagatgaggaaggaaggtatgtcgtgta 1127
QY 1052 acatggtcgaactcgtatcaagaataatgatacagatggtcgtcgtcgtcgtcgtcgtcgt 1111

DB 1128 acatggtcgaactcgtatcaagaataatgatacagatcgtcgtcgtcgtcgtcgtcgtcgtcgt 1187
QY 1112 agcacatcccatcaagaagc 1171
DB 1188 ccgcgcacacagatcaagaagc 1247
QY 1172 tccctcgtcgtgaggtcgc 1231
DB 1248 tgcgt 1307
QY 1232 ttatgtatgattcatcaacaacacagacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1291
DB 1308 tcatgtcgaactcgt 1367
QY 1292 gaaatctcgaagaagaaggttcaacaacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1351
DB 1368 gacagtgacgaagaagaaggttcaacaacactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1427
QY 1352 tccacactcgt 1411
DB 1428 tccacactcgt 1487
QY 1412 tcaaggtcgaaggtgaggtcgt 1461
DB 1488 tcgaggttgcaggttgcaggtcgt 1537

RESULT 4

US-09-738-626-836
Sequence 836, Application US/09738626
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIKAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 836
LENGTH: 1422
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-836

Query Match 40.7%; Score 594; DB 29; Length 1422;
Best Local Similarity 64.5%; Pred. No. 2.1e-154;
Matches 920; Conservative 0; Mismatches 500; Indels 6; Gaps 2;

QY 32 tcgagtagaattgcgcacatacaactcattcttcggtcgtgaaggtaactaccctg 91
DB 2 tggactcaaggttcgcgactcttaccatagcagggcagcgcgtccacagatctgtcgt 61
QY 92 ctgagaagaataatgcaggtctatagttcttcgtgagcggttattccgtcttaagcat 151
DB 62 cagaatattagatgcaggtctcatcaggttgcgaagggaattccgacagacgacct 121
QY 152 tgaagggttgcgaactctctgttccctccacatgacagttccacagacgcttccatcgt 211

QY 562 gtcttaacaaagaacaaaccactgcaacaagltgctgcccgaatgaacggtgttcc 621
 Db 1259 ggcttaacgtccgaacccacacaaagtacccgaagatgaagagagagctgctggagcttcc 1200
 QY 622 gaagagacaacaacaggtgtccacccgctctacagctcgaagggaggaacactctc 681
 Db 1199 gagagacacacacccggtgtcgaagagcttaccagatgacggagacggccctctc 1140
 QY 682 ttcccaagcacaacagtcgaagcgtgttacaagttccaaagttcgaatcgaatcgaagc 741
 Db 1139 ttcccccgcacatcaactcaacgactccgctcaccagagcaagtttgacaacctgtaaggt 1080
 QY 742 tgcgcacactccctatcgaatgaltcaaacggtgttcgaatgcatgcatgcaagcgaag 801
 Db 1079 tgcccccactctctccctgaatggtgtcgaatgagggctacacgattgattgctggcgaag 1020
 QY 802 acagctctgcatggttgaagcgtgtcgaagggcgtgcgaatccctcctcgtgc 861
 Db 1019 gttgccggtgtcgtggtatggtatggtgcaagggcgtgctgcttcaagcag 960
 QY 862 caagcgctcggtatcatcatcagaactcgaaccaatctgcgtctccaggtgcagt 921
 Db 959 gctggtgcccgtgtatgttactgagattgaccccatctgctccacggcccttatg 900
 QY 922 gaaggtacacaggtcgcgcgcatcgaggaagtcgtcaagatgtgatcttctgata 981
 Db 899 gaggggtccagaggtcctacatttgagagatgttctcgagagctgacatcttggtagcc 840
 QY 982 tgcacaggaactcgcatacatctctgttgaatgtgcccagatgaggaatgaagct 1041
 Db 839 accatggaacacaggaatcattatggtgacacatgaggaaggaaggaaggaagcc 780
 QY 1042 atgtcgtaacatcgccactgcatgaataagaaattgatacagatggtcctcgaat 1101
 Db 779 atgctttgcaacattgtgctactttgacaaatgattgacatgctgcttgagaccttac 720
 QY 1102 caaggtacacgaacatcccaatcaagcagaatagacatgtggaattcccaatgagc 1161
 Db 719 cctggtgtcaagccac 660
 QY 1162 cagcgt---atcctcctctgtgtgagggcgcgcctcttaaccttggtgcgtacagat 1218
 Db 659 aaacactggaatcattgtcttctgctgaggggtgtctcatgaaccttggtgcttactgac 600
 QY 1219 caaccatcttgcgtatgtcatgtcatcacaacacagacacacacacacacacacac 1278
 Db 599 caaccacagttgtgcatgctcctgctcatcacaacacacacacacacacacacacac 540
 QY 1279 t-----acgaagaagagagaatctcgaagaagaagtttacaacacttccgaacatc 1332
 Db 539 tggaaagagagagacactgtgcaatgacgaagaagaggtgtgcttctcccaagcactac 480
 QY 1333 gatgaagaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1392
 Db 479 gacgagaaaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 420
 QY 1393 cagaagcaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1452
 Db 419 aaagtcagaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1400
 QY 1453 cgttalia 1460
 Db 359 cggtacta 352

RESULT 6
 US-09-684-016-386668/c
 ; Sequence 386668, Application US/09684016
 ; GENERAL INFORMATION:
 ; APPLICANT: Royall, David K.
 ; APPLICANT: Liu, Jindong
 ; TITLE OF INVENTION: Annotated Plant Genes
 ; FILE REFERENCE: 38-21(15097)D

; CURRENT APPLICATION NUMBER: US/09/684,016
 ; CURRENT FILING DATE: 2000-10-10
 ; PRIOR APPLICATION NUMBER: US 09/554,617
 ; PRIOR FILING DATE: 2000-09-05
 ; NUMBER OF SEQ ID NOS: 463173
 ; SEQ ID NO 386668
 ; LENGTH: 1902
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-09-684-016-386668

Query Match 36.7%; Score 536; DB 27; Length 1902;
 Best Local Similarity 63.2%; Pred. No. 3.5e-138;
 Matches 915; Conservative 0; Mismatches 485; Indels 48; Gaps 4;

QY 34 gaatcagaatgtgcgaatcacaatcattcttcgcgcgcgaaggaatccttctgt 93
 Db 1772 gactacaaagtgaaagacacttccacagggcgaacttccgctcgaatgagcttgcgc 1713
 QY 94 gagaagaatgccaagttcttattgttcttctgtgagcgtatctcgtcttaagccatg 153
 Db 1712 gaggtcgaaggtccgggctcattgctgctgctgctgctgctgctgctgctgctgct 1653
 QY 154 aaggtgtcagaatctctgttctcctcacaatgaagtcgaagcagaacgcgtctcatcag 213
 Db 1652 aaggggcccggatctccgggtccctccacatgacacacacacacacacacacacacac 1593
 QY 214 acactcagctctgt 273
 Db 1592 accctcaccgccccttggcggcggaggtccgctggtgctcctcctcctcctcctcctc 1533
 QY 274 gatacagcgt 333
 Db 1532 gacacagcggccgcccgcac 1496
 QY 334 atccagctctcgtcgtgaaagggcgaacacacacacacacacacacacacacacacac 393
 Db 1497 -CGCGGTGTCGCTGGAAGGGGGAAGCCCTGAGAGATGCTGTGTCACCGAGGCGC 1440
 QY 394 gctcacaatgtccagatgtgtcaagggccacagcaggtgtcgtatgtgtgtgtgtgt 453
 Db 1439 tgcttgaactgtggggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1380
 QY 454 aaactcctatctcgaaggtctc-----gaattgaaagcgcgtgtgt 501
 Db 1379 acgctgctcattcagaagggcgtcgaagcccgagagagaggttcaagaagtcagagag 1320
 QY 502 ccagaagcacaagaagtcgaac 561
 Db 1319 cccgacccggaggtccac 1260
 QY 562 gtcttaacaaagaacaaaccactgcaacaagltgctgcccgaatgaacggtgttcc 621
 Db 1259 ggcttaacgtccgaacccacacaaagtacccgaagatgaagagagagctgctggagcttcc 1200
 QY 622 gaagagacaacaacaggtgtccacccgctctacagctcgaagggaggaacactctc 681
 Db 1199 gagagacacacacccggtgtcgaagagcttaccagatgacggagacggccctctc 1140
 QY 682 ttcccaagcacaacagtcgaagcgtgttacaagttccaaagttcgaatcgaatcgaagc 741
 Db 1139 ttcccccgcacatcaactcaacgactccgctcaccagagcaagtttgacaacctgtaaggt 1080
 QY 742 tgcgcacactccctatcgaatgaltcaaacggtgttcgaatgcatgcatgcaagcgaag 801
 Db 1079 tgcccccactctctccctgaatggtgtcgaatgagggctacacgattgattgctggcgaag 1020
 QY 802 acagctctgcatggttgaagcgtgtcgaagggcgtgcgaatccctcctcgtgc 861
 Db 1019 gttgccggtgtcgtggtatggtatggtgcaagggcgtgctgcttcaagcag 960
 QY 862 caagcgctcggtatcatcatcagaactcgaaccaatctgcgtctcgaagctgcagt 921

QY 1169 tccctctctgtgagggccgctcttaacctgtgctgctacaggtccacctctt 1228
 Db 1196 tcaatgtgtgtcggagggcggtgtgtgaaacctgtgaaatgcaacggcgacccctgt 1255
 QY 1229 tctgtgtcaatgtgtctcaaacagagacacgtcgcctcaactcgtacactcaagaaga 1288
 Db 1256 tctgtgtgagcaacagcttctgcaaacagagacacgtcgcctcaactcgtgtgtgaaaga 1315
 QY 1289 gagggaatctcgagaaaggtttacacacttccgaagacatctcgtatgaagagtcgtc 1348
 Db 1316 agcagcagatcgacaaacagaggtgtacccgctgcccgaacactcgtacgagaagtgctc 1375
 QY 1349 gctccactcgtatctcgtatgtccacttcaaaagcttacaagaagcagctgtact 1408
 Db 1376 gatcatgtcgaagccctcgtcgtcgtacacgaagctgacccaagagagcagcggaat 1435
 QY 1409 acataacgcttcaggtgtgaggtctcttacaagctcgtatgtctaccgtatataa 1461
 Db 1436 acctcgcgtcgaagcagaggtcctctacaagcgcgaacactacgctactga 1488

RESULT 8

US-09-703-708-7943
 ; Sequence 7943, Application US/09703708
 ; GENERAL INFORMATION:
 ; APPLICANT: Bower, Stanley G.
 ; APPLICANT: Hinkle, Gregory J.
 ; TITLE OF INVENTION: Xanthomonas campestris Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15804)C
 ; CURRENT APPLICATION NUMBER: US/09/703,708
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: US 60/164,320
 ; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: US 60/183,791
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 18992
 ; SEQ ID NO 7943
 ; LENGTH: 1434
 ; TYPE: DNA
 ; ORGANISM: Xanthomonas campestris
 ; US-09-703-708-7943

Query Match 36.4%; Score 531.6; DB 28; Length 1434;
 Best Local Similarity 62.3%; Pred. No. 5.3e-137;
 Matches 909; Conservative 0; Mismatches 509; Indels 42; Gaps 3;

QY 14 caactactgtgtcctcattcgaagatgacgaagatgacgaatcaactcactgttctcgcc 73
 Db 5 caaaaacggtccacacacacgattacacagatcgcgcgaacatcctctgcccgaattgggccc 64
 QY 74 gtaaggaacttaccctgtgtgagaagaatgacaggtctctatgttctctgtgagcgtt 133
 Db 65 gtaaggaactgtgacatcgcgaagacagatgacgaggtcgtgtgacatgctccgcaagc 124
 QY 134 attcgcctcttaagcatttgaaggtgttcgaatctctgtgttctccatcgaatgacgtc 193
 Db 125 atgacgagacacacacgctggaagacgctccgacacacgctcgtcgtgacatgacatcc 184
 QY 194 agcagcgcgttccatcgaagacacacgactcgtgtgtgtgtgtgtgtgtgtgtgtgt 253
 Db 185 agacccgctgt 244
 QY 254 gcaaatctctcttcaacagatacagcgcgtcgtctgtctgtgtgtgtgtgtgtgtgtgt 313
 Db 245 gcaaatctctcttcaacagatacagcgcgtcgtctgtctgtgtgtgtgtgtgtgtgtgt 289
 QY 314 caccagagaagcagcgcgttatccagcttctgctggaagggcgaacactccagaaat 373
 Db 290 caccagagaagcagcgcgttatccagcttctgctggaagggcgaacactccagaaat 337
 QY 374 actggaagaacacatacagcgcgtctcactgacagatgtgtca-----agccc 421

Db 338 attggactgaacccctggaacgcgtgacttcaacctgcccgaacgcgtacacggcc 397
 QY 422 cacaagcagctgtgtgatgt 481
 Db 398 cgaagcgt 457
 QY 482 tcgaacagcgcgt 541
 Db 458 tcgaacagcgcgt 514
 QY 542 tcttctgaacactcaagaaggtcttcaacacagaagaacacgaatgtgtgtgtgtgtgt 601
 Db 515 tcaagcagctgtcgaagcgt 574
 QY 602 ccgagcagcagcgt 661
 Db 575 aggaactggaagggcgt 634
 QY 662 agaaaggagggcaacactcctctccagacatcaacgctgaacgctgttacaagaatcca 721
 Db 635 ccgaagcgcgcgaacactcgt 694
 QY 722 agtctgaatacatctacagcgtgcgcgcacacacacacacacacacacacacacacac 781
 Db 695 agtctgaatacatctacagcgtgcgcgcacacacacacacacacacacacacacacac 754
 QY 782 atgtcatgtacgc 841
 Db 755 acgtgtatgtcgc 814
 QY 842 ggcgtcaataccctcgc 901
 Db 815 ggc 874
 QY 902 ggcgtctcagcgtgc 961
 Db 875 ggc 934
 QY 962 atgtcgtatctctcgtgtacatgacagaggaacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1021
 Db 935 ggc 994
 QY 1022 cccaagatgaagataagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1081
 Db 995 aggcgtatgaagataagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1054
 QY 1082 cagatgctctcattgaatacccaagcgtacgaacacacacacacacacacacacacacac 1141
 Db 1055 tcatgtcgt 1114
 QY 1142 tgtgtggaatcccaagatgac 1201
 Db 1115 agtatgt 1174
 QY 1202 ttgctcgt 1261
 Db 1175 ttgctcgt 1234
 QY 1262 tctgtcagcgtcgt 1321
 Db 1235 ttgctcagcgtcgt 1294
 QY 1322 cgaagcgtcgt 1381
 Db 1295 cgaagcgtcgt 1354
 QY 1382 caaagcgttacaagaagcagcgt 1441
 Db 1355 caaagcgttacaagaagcagcgt 1414
 QY 1442 ctgtatgtcgt 1461
 Db 1415 cgtgtatgtcgt 1434

QY 14 caactactgtgtcctcattcagagatgacagaaatgacacactcactcattgtctcggcc 73
 DB 5 caaaaacccgttccacacacagattacaaagatcgccagacatccctcgtcgtttggggcc 64
 QY 74 gtaaggaacttaccctgtctgtaaggaagaaatgacaggtcttatgttctcgttggagct 133
 DB 65 gcaaggaagctcgagacatcgccgagacagagatcgccggtctgttctgacatcgccggaagc 124
 QY 134 attcgccttctaagccattgaagaggtgtcagaatctcgtgtcctcctccacagacagctcc 193
 DB 125 atgacgagaccagcagcgtctgaagaagcgtctcgcacatccacgcgcgtctgacacagacatcc 184
 QY 194 agacacacgtctcctcactcgaagacactcacaagctctgtgtgtatgacatgaggtctcct 253
 DB 185 agaccccggtgtctgatacgaagacgtcgaagacatcgccgaatgtgctgctggtcctcgt 244
 QY 254 gcaacatctctctacacacagatacagccgctgtcgtctatcgttctcggcccaacagaga 313
 DB 245 gcaacatctctctacacacagatacagccgctgtcgtctatcgttctcggcccaacagaga 289
 QY 314 caccagagaagacagccggtatcccaatctcgtcgtgaagagcgaaacacacacagat 373
 DB 290 -----ccaccggaacagcgtgtgttctcgtcgtgaagagcgaaacacacagat 337
 QY 374 actggaagaaacataacccgctctcactcactcactcactcactcactcactcactcactc 421
 DB 338 atggagactcgcacactcgcgtgtgacgtctgacactcactcactcactcactcactcactc 397
 QY 422 cacaagaggtgtgacgt 481
 DB 398 cggagagctgt 457
 QY 482 tcgaaacagccggt 541
 DB 458 tcgaaacagccggt 514
 QY 542 tctctgtcactcactcactcactcactcactcactcactcactcactcactcactcactcactc 601
 DB 515 tcaagagcactcactcactcactcactcactcactcactcactcactcactcactcactcactc 574
 QY 602 ccggaatgacaggt 661
 DB 575 aggactggaagagcggtctcgtcgtgaagagacacacacacacacacacacacacacacac 634
 QY 662 agaagagagggcacaactcctctcctccagacatcaacgtcaacacacacacacacacacac 721
 DB 635 ccgaagccggaacactcgt 694
 QY 722 agtgcgataacatctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 781
 DB 695 agtgcgataacatctcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 754
 QY 782 atgtcagatcgc 841
 DB 755 acgtgatgtcgc 814
 QY 842 ggcgcacatcctcctcgtcgtcgc 901
 DB 815 ggcgcgcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 874
 QY 902 ggcgcctcagcgtcgc 961
 DB 875 ggcgcctcagcgtcgc 934
 QY 962 atgtcagatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1021
 DB 935 ggcgcgcacatctatgtcgc 994
 QY 1022 cccagatgaaagaaagcgtatgt 1081
 DB 995 aggcgatgaaagaaagcgtatgt 1054

QY 1082 cagatggtcctcagatgaataccagcagatcaagcacaacccacacacacacacacacacacac 1141
 DB 1055 tcgatgctcgtatgacacttaagagcgtgtggaagatgaatgaatgaatgaatgaatgaatgaatga 1114
 QY 1142 tgtggaatcccaatggtcagcagcgtatccctcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgt 1201
 DB 1115 agtatgtgtcgtcagcagcgtatccctcctcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1174
 QY 1202 ttggtcgtcgtcagcagcgtatccctcctcctcctcctcctcctcctcctcctcctcctcctcct 1261
 DB 1175 tggcgtcgtcgtcagcagcgtatccctcctcctcctcctcctcctcctcctcctcctcctcctcct 1234
 QY 1262 tcgtcagcgtcagcgtcagcgtatccctcctcctcctcctcctcctcctcctcctcctcctcctcct 1321
 DB 1235 tggcgtcagcgtcagcgtcgtgtggaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1294
 QY 1322 cgaagcactcgtatgaagaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1381
 DB 1295 ccaagcactcgtatgaagaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1354
 QY 1382 caagcgtcagcagcagcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1441
 DB 1355 ccacccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1414
 QY 1442 ctgattcgtcagcgtatga 1461
 DB 1415 cgtatcattacgcgtactaa 1434

RESULT 11
 US-09-703-708-691/C
 ; Sequence 691, Application US/09703708
 ; GENERAL INFORMATION:
 ; APPLICANT: Bower, Stanley G.
 ; APPLICANT: Hinkle, Gregory J.
 ; TITLE OF INVENTION: Xanthomonas campestris Genome Sequences and Uses Ther
 ; FILE REFERENCE: 38-10(15804)C
 ; CURRENT FILING DATE: US/09/703,708
 ; PRIOR APPLICATION NUMBER: 2000-11-02
 ; PRIOR FILING DATE: US 60/164,320
 ; PRIOR APPLICATION NUMBER: 1999-11-10 US 60/183,791
 ; PRIOR FILING DATE: 2000-02-22
 ; NUMBER OF SEQ ID NOS: 18992
 ; SEQ ID NO 691
 ; LENGTH: 19798
 ; TYPE: DNA
 ; ORGANISM: Xanthomonas campestris
 US-09-703-708-691

Query Match 36.4%; Score 531.6; DB 28; Length 19798;
 Best Local Similarity 62.3%; Pred. No. 1.3e-136;
 Matches 909; Conservative 0; Mismatches 509; Indels 42; Gaps 3;
 QY 14 caactactgtgtcctcattcagagatgacagaaatgacacactcactcattgtctcggcc 73
 DB 16248 CAAAAACGGTTCCACACACGATTACAGAGATCGCGACATATCCCGGGCGATTGGGCC 16189
 QY 74 gtaaggaacttaccctgtctgtaaggaagaaatgacaggtcttatgttctcgttggagcgt 133
 DB 16188 GCAAGGAGCTGGACATCGCCGACGACGAGATCGCGGCTGATGTGATTCGATCCGCCGCAAGC 16129
 QY 134 attcgccttctaagccattgaagaggtgtcagaatctcgtgtcctcctcacaatgacagctcc 193
 DB 16128 ATGCGGACGACCAAGCGCCTGAAAGACGTCGCCATACCGGCTCGTCACATGACCATCC 16069
 QY 194 agacacgcgtcctcactcagacacactcacaagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 253
 DB 16068 AGACCGCGGCTGATGATGAAGCTCAAGGACATCGCGCCCAATGTGCGCTGGGCTCGT 16009
 QY 254 gcaacatctctcactacacagatacagccggtgtcgtcgtatcgttctcggcccaacagaga 313


```

Db 15678 AGACATCGAAGGCGCTTCTGTAAGACACACCGGCGCTGACCGCTGTACCGATGTG 15619
QY 662 aagaaggaggcaaacctcctctcccaagcatcaacgcaacgctgttacaagatcca 721
Db 15618 CCGAAGCCGCGCAACTCTGATCCCGGCATCAAGCTCAAGACATCCGTCACCAAGAGCA 15559
QY 722 attcgataacatacctaagagctgcgcgcacccctcctatcgaatgtatcaacgctgtccg 781
Db 15558 ACTTCACCAACTCTACGGGTGCGCGCATGCTGCGCATGCGCTGACGCGCATGCG 15499
QY 782 atgtcatatcgcggaagaacagctctcgtcatggtgtttagcgagatgtcggaaaggtc 841
Db 15498 AGGTGATGCTGCGCGGCAAGGTCCGCGTGTGCTGCGGTACGCGCATGTCGGCAAGGCA 15439
QY 842 ggcgtcaatccctcgtggaagaagcgctcgtatatacatcaacgaacacccaatctc 901
Db 15438 GCGCGCCGCTGCGTGGCGTACGCGTACGCGCGCTGATGTCACCGCATGTCGACCCGATCT 15379
QY 902 ggcgtctcgaagctgcgaatggaagctacacagctcgcgcgaatcgaggaagctcgaag 961
Db 15378 GCGCGCTGAGGGGTGATGGAAGGCTTCGAGTCAACACCATCGAATCCACCTGGGCG 15319
QY 962 atgtcatatctcgttatacatgcaagaagaatcgatatacgtctgtttagatagtg 1021
Db 15318 GCGCGCATCTATGTATGTCACCAACCGGCAACAGGACATCATCAACCGTCGACCTGCG 15259
QY 1022 cccagatgaaggaataagctatgttcgttaacatcgccacttcgaataaagatgata 1081
Db 15258 AGGCATGAAGACACAGGCGCATGCTGCAACATCGGCGCATCTGCAACAGGATCCAGG 15199
QY 1082 cagatgagctctgaataacacagctcgaacgaatcccaatcaagcagaatacagaa 1141
Db 15198 TCGATGCGCTGATGCACTTAAGGCGCTGGAAGATCAACATCAAGCGCGAGTGCACA 15139
QY 1142 tgtggaatcccaagatggcgaacgtatccctctctctgttgaagcgcccttcaacc 1201
Db 15138 AGTATGTGTTGCGCAACGGTAACGGCATCTGCTGCGCGCGCGCGCGCTGTGTAAC 15079
QY 1202 ttggtgtgcgtacaggtacacccatcttctgtatgtcaatgtcaltacaaacagagac 1261
Db 15078 TGGGTGCGCCCAACCGGCGCACCGCTGCTGATGTCACACCTGCGCAACCAAGACCC 15019
QY 1262 tcgctcagctcgaacgtcgaagaagaagagaaatctcgaagaagaagtttacaatctc 1321
Db 15018 TGGCGCATGTCGACTGTGGGAAAGCGGACACGCTACGAAGAAGAGGCTTACATCTCGC 14959
QY 1322 cgaagcatctcgaatgaagaagctcgtcgtccctcgaatcgtcgtatcgtatgcaccta 1381
Db 14958 CCAAGCACCCTGATGTAAGAAAGTGGCCCGCTGCACCTGGAAGAAGTCCGCGCTCAAGCTGA 14899
QY 1382 caaagctaacagaagaagctgtaactacatcaacgltccagttgaaggttccttaag 1441
Db 14898 CCACCCCTACCAAGGACCGGCGGCTGCTGCGGTGAGAGTGGCTGGCGCGTCAAGC 14839
QY 1442 ctgattcctacgttatiaa 1461
Db 14838 CGGATCATTTACCGCTACTAA 14819

```

```

RESULT 13
US-60-183-791-691/C
; Sequence 691, Application US/60183791
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; TITLE OR INVENTION: Xanthomonas campestris genome sequences and uses thereof
; FILE REFERENCE: 38-10(1584)B
; CURRENT APPLICATION NUMBER: US/60/183,791
; CURRENT FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 18992
; SEQ ID NO 691
; LENGTH: 19798

```

```

; TYPE: DNA
; ORGANISM: Xanthomonas campestris
US-60-183-791-691

Query Match 36.4%; Score 531.6; DB 57; Length 19798;
Best Local Similarity 62.3%; Pred. No. 1.3e-136;
Matches 909; Conservative 0; Mismatches 509; Indels 42; Gaps 3;

QY 14 cactactgtgtcctcaatgagatagaagaatgcccgaatcaacatcctcattcctgcgc 73
Db 16248 CAAAAACCGTTCACACACCATTAAGATGCGCGACATCTCCGCGCGATTTGGGCGC 16189
QY 74 gtaagaaactaacctctgtcgaagaagaatgcaaggtttatgttctcgtgagaggt 133
Db 16188 GCAAGGAGCGTGGACATGCGCGGACGACGAGATGCGCGCTGATGATCGCGCGAAGC 16129
QY 134 attcgccttaagcaatggaaggtgtcgaatctgttccctcacaatgacagttcc 193
Db 16128 ATGCGGACGACCAAGCCGCTGAAGAAGATGCGGATCAGCGGCTGCGATGACATGCCATCC 16069
QY 194 aagacagcgctcctcatagagacatcaagctctgtgtcgtatcgttgcgtcgaaggtcctc 253
Db 16068 AGACCGCGTGTGATGAAAGCTCAAGGACATGCGCGCAATGCGCTGCGCGCTCGT 16009
QY 254 gcaacatctctctacaaagatacaagacgctgtcgtatcgttgcgtcgaaggtcga 313
Db 16008 GCMAACATCTCTCCACCGACGACGACGCGCGCGCGCGCTGCGCGCTGCGCGCTGCG 15962
QY 314 caccagaagaagcagcggtatcccaagctcgtcgtcgtgaaggtcgaagaacatccagaat 373
Db 15963 -----CCACCGGCGACGCGCGGTGCTGCGTGAAGGCGGACGCTGGAAGAT 15916
QY 374 actgagagaacacatacaccgctcctcacaatgcaagatgta-----agggc 421
Db 15915 ATTGGGACATCCACTTGAGACGCGCTGACCTTACCTTGCCCGACGCGACCTGACCGGCGC 15856
QY 422 caaagcaggtgtcgtatgaggtgtgtatgtctacatcctcctcacaaggtcgtcgaat 481
Db 15855 CCGAGCTGTGTGTGAGAGAGCGGCGGCGGACGTCACCTTGATTCACAAAGGCTTATGAGC 15796
QY 482 tgaagaacagcggtgtgttcgaagagcgaagaagctgacaacctcgaatacagctgcg 541
Db 15795 TCGAAACCGGACACCTGCGGTGTCACGACGCGG-----CTGCTGCGCAAGAAAGGCGTCA 15739
QY 542 tctctcacaacccaagcaggtcttcaaccaagaagaacccaatggaacaaagttgtcgt 601
Db 15738 TCAAGGCACTGCTCAACCGCGGTGCTGTGAGCGCTCCGGGTACTGTGGCGCGGTGTCA 15679
QY 602 ccggaatgaacggtgttccgaagaagaacaacaggtgtccacccgctcctacacgtcgtc 661
Db 15678 AGGACTGGAAGGCGTCTGTGAAGAGACACACCGCGGTGAGACCGCGCTGTACCAATTTG 15619
QY 662 aagaaggaggcaaacctcctctcccaagcatcaacgtaacgaagctgttacaagttcca 721
Db 15618 CCGAAGCCGCGCAACTCTGATCCCGGCATCAAGCTCAAGACATCCGTCACCAAGAGCA 15559
QY 722 attcgataacatacctaagagctgcgcgcacccctcctatcgaatgtatcaacggtgtccg 781
Db 15558 AGTTGACCAACCTCTACGGGTGCGCGCATGCTGCGCATGCGCTGCAAGCGCGCATGCG 15499
QY 782 atgtcatatcgcggaagaacagctcgtcatggtgtttagcgagatgtcggaaaggtc 841
Db 15498 AGGTGATGCTGCGCGGCAAGGTCCGCGTGTGCTGCGGTACGCGCATGTCGGCAAGGCA 15439
QY 842 ggcgtcaatccctcgtggaagaagcgctcgtatatacatcaacgaacacccaatctc 901
Db 15438 GCGCGCGCTGCGTGGCGTACGCGTACGCGCGCTGATGTCACCGCATGTCGACCCGATCT 15379
QY 902 ggcgtctcgaagctgcgaatggaagctacacagctcgcgcgaatcgaggaagctcgaag 961
Db 15378 GCGCGCTGAGGGGTGATGGAAGGCTTCGAGTCAACACCATGGAATTCACCGCTGGGCGC 15319

```

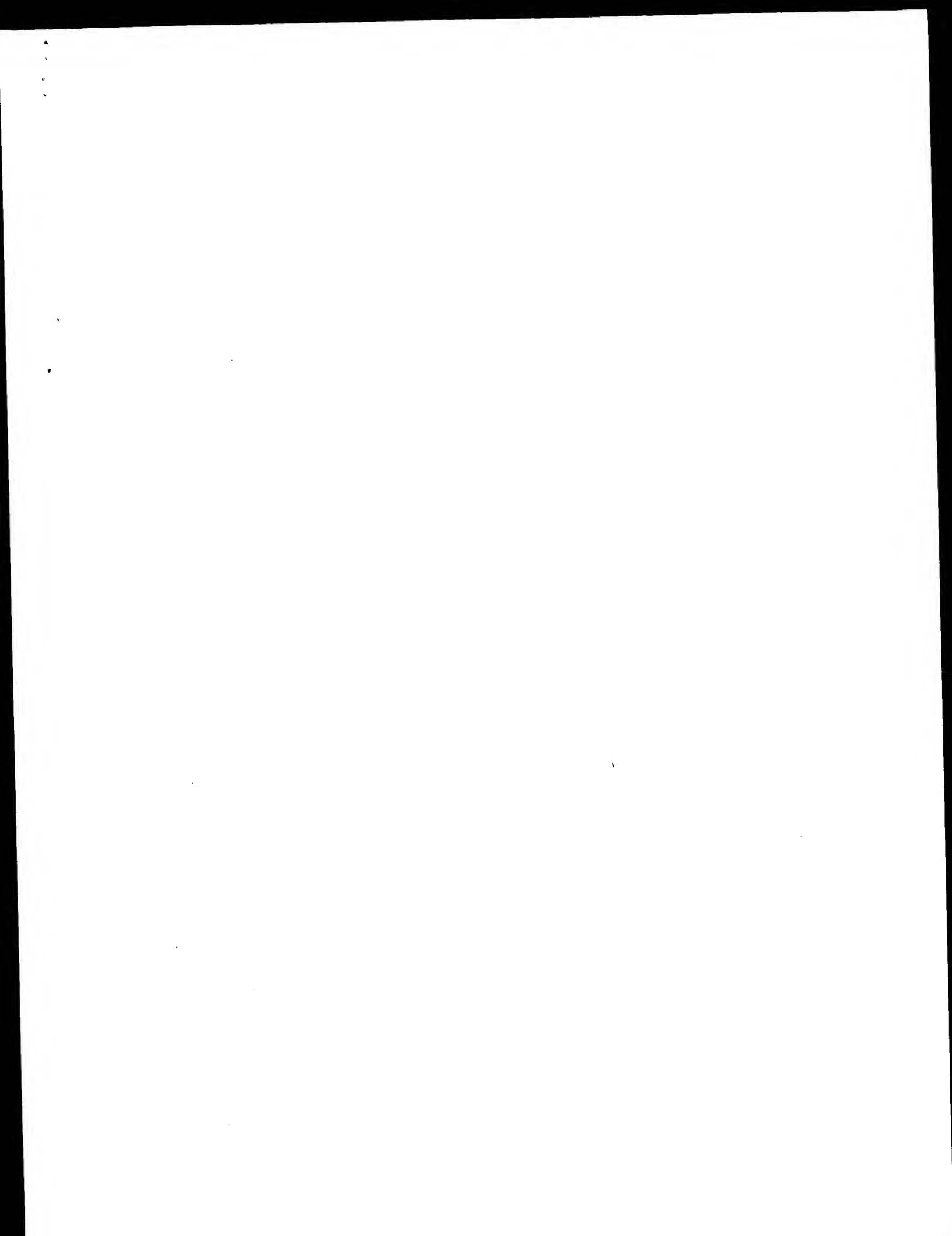
QY 962 atgtcatatcttgcgtacatgacagaagaactgsgatatactctgttgcacatgttg 1021
Db 15138 GCGGCGCATATGTACACACCGCAACAGACATCATACCGTGGAGACCTGC 15259
QY 1022 ccagatgaagataagatctatgttgcgaacacgcgcacatcgcgaataatgata 1081
Db 15258 AGGGATGAAAGACAGGCGATCTGTGCAACATCGGCCACTTCGACACAGATTCAGG 15199
QY 1082 cagatgacctcatgaatacccaagcgaataagacatcccaatcaagccgaataagaca 1141
Db 15198 TCGATGCCCTGAATCACTTAAGGCGTGAGAAATCAATCAATCAACCGAGTGCACA 15139
QY 1142 tgtggaattccagatggccacgcatctcctcttcttgaaggcgcccttcaacc 1201
Db 15138 AGTATGTTTGGCAACGGTAACGCATCTCTCTGCTGGCGAGCGCGCTGTGAACC 15079
QY 1202 ttgctgcgtacaggtaccccatcttctgttatgtcaatgttaatacaaacagacac 1261
Db 15078 TGGCTGCGCCACCGGCCACCGAGCTGTGATGTCAACTGCTTGCACCAACAGACCC 15019
QY 1262 tcgtcagctgcgactctacgaagaagaagaatctcgaagaagaagtttacacacttc 1321
Db 15018 TGGCCAGATCGACCTGTGGAAAGCGGACAGCTACGAGAGAAAGGTCTACATCTGTC 14959
QY 1322 cgaagcatctcgatgaagaagtcgtcgcctccacactcgatctcgatgtccacctta 1381
Db 14958 CCAACACCTGTGATGAAGAATGGCCCGCTGCACCTGGAAGATCGCGCTCAAGTGA 14899
QY 1382 caaagcttacacagaagcgtgactacatcaacggttcagltgaggttcttacaagt 1441
Db 14898 CCACCTTCACCAAGGACCGGCGGACTACTCGCGTGAGAGCTGGCCCGTAACAAGC 14839
QY 1442 ctgatctaccggttaata 1461
Db 14838 CGGATCATTAACCGCTACTAA 14819

RESULT 14
US-09-606-740A-223
Sequence 223. Application US/09606740A
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
METABOLIC PATHWAY PROTEINS
FILE REFERENCE: BG1-121CP
CURRENT APPLICATION NUMBER: US/09/606,740A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142,101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148,613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187,970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931415.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931418.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: DE 19931434.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931435.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931443.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931453.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931457.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931465.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931478.0
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931510.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931541.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931573.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931592.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931632.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931634.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932130.2
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932186.8
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932206.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932227.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932228.7
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932229.5
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932230.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932922.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932926.5
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932928.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933004.2
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19940764.9
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940765.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940766.5
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940832.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941380.0
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941394.0

Mon Aug 19 15:16:25 2002

us-09-759-990-1.rnp



9 AspPheIyValAlaIAspleuSerIeuAlaIaPheIyArgIySgluI 25
 84 TACCTTGGTGGAGAGAAATGCCAGGCTTATGGTTCTTGGTGAGCGTT 133
 25 ePrlIeuAlaIgluHisIgluMetProGlyIleuMetAlaIleArgIySgluT 42
 134 ATTCGCGCTTAAAGCCATGTGAAGGTCATGAAATCTCTGGTTCCTCCAC 183
 42 yAlaIgluAlaIgluProIeuAlaIgluAlaArgValThrGlySerIeuHis 58
 184 ATGACAGTCCAGACAGCGCTCCATGAGACACTCACAGCTCTTGGTGC 233
 59 MetThrValIgluThrAlaValIleuIleuIleuThrIleuValAlaIleuGlyAl 75
 234 TGATGTGAGATGGCTTCCTGCAACATCTCTTACACAGATACAGCGG 283
 75 aGluValATGTTPrAlaSerCysAsnIlePheSerThrIleAspHisAla 92
 284 CTGCTGTATCGTTGTCGCCCAACAGGACACAGAGAGAGCCCGGT 333
 92 IAlaIAlaIleIleAlaValIgluProAsnGlyThrProAspAsnProGluI 108
 334 ATCCCACTCTTCCGCTGGAAGGCGAAACACTCCGAAATATCGGAGAA 383
 109 ValProValIleAlaIleATrIySgluIleuIleuIleuIleuIleuIleu 125
 384 CAGATACCGGCTCTCATGTCGACAGATGTCAA.....GGCCACAGC 427
 125 sTrIgluIleIleAlaIleuThrTrpProAsnThrProThrIleGlyProAsnM 142
 428 AGGTTGTGATGATGTGGTATGCTACACTCTCTCATCTCCAGGCGTTC 477
 142 eTlleuAspAspGlyGlyAspAlaThrIleuValHisIleuSgluVal 158
 478 GAATTCGAACAGCGCGCTGCTCCAGAGCCAAAGAAAGAGAGAGAGAG 527
 159 GluTrpIleuIyAspGlyIyValProSerValAspThrAlaGluSerAs 175
 528 CGAATACCGGCTCTCTGCTTACACTCAAGAGGCTTCAACCAAGACA 577
 175 pGluHisArgValIleIleuGluIleuIleuThrArgThrValIgluIleu 192
 578 AGAACCACTGGACACAGTGTGCGGATGAGAGAGGTTTCCGAAAG 627
 192 roGluIyStrpThrIleuAlaSerGluIleArgIyValThrIgluIleu 208
 628 ACAACACAGGTCACCGGCTTACAGCTCGAGAGAGAGAGAGAGAG 677
 209 ThrThrThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 225
 678 CTTCTTCCAGCCATCAACGTCACAGAGCTGTTCACAAAGTCCAACTTCG 727
 225 uIleuPheProAlaIleAsnValAsnAspAlaValThrIySerIyPheA 242
 728 ATACATCTAGGCTGCGGCGCACTCCATTATGATGGTATCAACGCTGCT 777
 242 spAsnIyStrpGlyCysArgHisSerIleuIleAspGlyIleAsnArgAla 258
 778 TCCCATCTCATGCGGCGGCAAGACACTCTCGATGGTATTCAGCGCA 827
 259 ThrAspAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 275
 828 TGTGCGCAAGGCTGCGCTCAATCCCTCGTGGCCAAAGGCGCTGCGCTTA 877
 275 pValGlyIySgluCysAlaIleuSerIeuArgSgluIleuIleuIleuValI 292
 878 TCATCACAGAACTGACCCATCTGCGCTCCAGGCTCCATGGAAGG 927
 292 IeIleuThrIleuIleuProIleuIleuIleuIleuIleuIleuIleuIleu 308
 928 TACCAAGTCCGCGCATGAGAGAGAGTGGTCAAGATGATGATCTTCGT 977
 309 PheIleuValIleuThrIleuAspIleuValIleuAspIyAlaAspIleuPhe 325

978 TACATGCACAGAAACTGCAATATCATCTGTGTGACATGATGCCACGA 1027
 325 IThrThrThrIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 342
 1028 TGAAGATGAGCTATTGTGCGTAAACATCGCCACTTCATTAAGAAAT 1077
 342 eIyHisIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 358
 1078 GATACAGATGGCTTCATGAATACCCAGGATCAACACATCCCAATCAA 1127
 359 AspMetAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 375
 1128 GCCAATATACGATGTGGGAATTCACAGATGCGCAAGCTATCCCTTC 1177
 375 sProGluValHisStrpThrTrpThrTrpProAspGlyIleuIleuIleuVal 392
 1178 TTGCTGAGGCGCGCTTCTTAACCTTGGCTGGCTACAGTCAACCATCT 1227
 392 euserGluGlyArgIleuIleuIleuIleuIleuIleuIleuIleuIleu 408
 1228 TTGCTTATGATGATGTCATTCACAAACAGACACTGCTCAGCTGACCT 1277
 409 PheValIleuSerAsnSerPheAlaAspGluIleuIleuIleuIleuIleu 425
 1278 CTACGAAAGAGAGAGAAATCTCGAAGAGGTTTACACATTCGGAAGC 1327
 425 uPheThrIyStrpProAspGlyIleuIleuIleuIleuIleuIleuIleu 442
 1328 ATCTCGATGAAGAACTGCTGCTCCAGCTCGAGATCTCGATGTCAC 1377
 442 IleuAspGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 458
 1378 CTTACAAAGCTTACACAGAGAGGCTGACTACATCAACCTTCAGTTGA 1427
 459 IeuthrThrIleuArgProGluGluIleuIleuIleuIleuIleuIleu 475
 1428 GGGTCTTACAAAGTGTGCTTACCGTTAT 1458
 475 uGlyProTrpIyAspIleuIleuIleuIleuIleuIleuIleuIleuIleu 485

seq_name: sp_bacteria:Q936D6

seq_documentation_block: PRT: 482 AA.
 ID Q936D6; PRELIMINARY;
 AC Q936D6;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ADHICASE.
 GN MTM.
 OS Streptomyces argillaceus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=41951;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fernandez Lozano M., Lombo F., Brana A.F., Sales J.A., Mendez C.;
 RT "Generation of a Streptomyces argillaceus mutant by deleting two
 RT methyl cycle-related genes located within the mlbA-mycA gene cluster
 RT leads to higher production of blibramycin.";
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBD databases.
 DR EMBL; AJ416377; CAC94890.1;
 SQ SQUENCE 482 AA; 52563 MW; B3C4D7ADB25606 CRC64;

alignment_scores:

Quality: 1625.50 Length: 477
 Ratio: 4.064 Gaps: 2
 Percent Similarity: 83.857 Percent Identity: 65.828

alignment_block:

US-09-759-990-1 x Q936D6 ..

Align seg 1/1 to: Q936D6 from: 1 to: 482

```

34 GAGTACAGAAATGGCCGACATCCATCCATGTTCGGCCGTAAGAACT 83
   : : : : : : : : : : : : : : : : : : : : : : : : : :
7  AspPheIysValAlaAspLeuSerLeuAlaGlnPheGlyArgGlyGluI 23
84 TACCTTCGTCGAGAGAAATGCCAGGCTTATGGTCTTCGTCGAGCTT 133
   : : : : : : : : : : : : : : : : : : : : : : : : : :
23 eThrLeuAlaGlnHisGlnMetProGlyLeuMetSerIleArgArgGlu 40
134 ATTCGCGTTCAGAGCATTTGAAGGTCGACATCTCTGTCGTCGTCGAC 183
   : : : : : : : : : : : : : : : : : : : : : : : : : :
40 yAlaGlnAlaGlnProLeuAlaGlyAlaArgValThrGlySerLeuHis 56
184 ATGACAGTCCAGACAGCGCTCCATGAGACACTCAGACGCTTCGTGTC 233
   : : : : : : : : : : : : : : : : : : : : : : : : : :
57 MetThrValGlnThrAlaValLeuIleGlnThrLeuValAlaLeuGly 73
234 TGATGTCAGATGGGCTTCCTGCAACATCTCTACACAGATACAGCCG 283
   : : : : : : : : : : : : : : : : : : : : : : : : : :
73 aGlnValArgTyrPalaSerCysAsnIlePheSerThrGlnHisAla 90
284 CTGCTGCTATGTTGTCGCCACAGAGGACACACAGAGGACGCGCT 333
   : : : : : : : : : : : : : : : : : : : : : : : : : :
90 lAlaIleAlaIleValGlyProAspGlyThrProAspAsnProArgAla 106
334 ATCCAGTCTTCGCTGGAAGGCGAAACATCCAGAAATACGGAGAA 383
   : : : : : : : : : : : : : : : : : : : : : : : : : :
107 ValProValPheAlaTyrPylsGlyGlnThrLeuGlnGlyTyrPyl 123
384 CACATGCCCGGCTTCACATGCCAGATGGCA.....GCCGACAGC 427
   : : : : : : : : : : : : : : : : : : : : : : : : : :
123 sThrGlnGlnAlaLeuThrTyrProAspSerProThrGlyIleProAsn 140
428 AGCTGTGTCATGATGGTGTGATGCTACACTCTCATCTCCAAAGGCTTC 477
   : : : : : : : : : : : : : : : : : : : : : : : : : :
140 eTlleuAspAspArgGlyAspAlaThrLeuLeuValHisGlyGlyAl 156
478 GAATTCGAAGACGCGCTGCTTCAGAGCAACAGAAAGCTGACAACT 527
   : : : : : : : : : : : : : : : : : : : : : : : : : :
157 GlnTyrGlnLysAspGlyLysValProSerValAspThrAlaGlnAsn 173
528 CGAATTCGCGCTTCCTACACTCAGAGGCTTCACAAAGCAAGCA 577
   : : : : : : : : : : : : : : : : : : : : : : : : : :
173 pGlnHisAlaGlyAlaIleLeuGlnLeuLeuAsnArgThrIleThrAsp 190
578 AGAACCACTGGACACAGTGTGTCGCCGATGAACGCTGTTCCGAAG 627
   : : : : : : : : : : : : : : : : : : : : : : : : : :
190 eTlleuHisAlaGlyAlaIleLeuAlaSerGlnIleArgGlyValThrGln 206
628 ACAACAAAGGTCGTCGCGCTTCACAGCTCAGAGAGAGGCGCAACT 677
   : : : : : : : : : : : : : : : : : : : : : : : : : :
207 ThrThrThrGlyValHisAlaGlyLeuTyrGlnMetGlnArgAspGly 223
678 CTTTCCAGCAATCAGCTCAAGCGCTGTTTACAAAGCCCAAGTCCG 727
   : : : : : : : : : : : : : : : : : : : : : : : : : :
223 uLeuPheProAlaIleAsnValAsnAspAlaValThrLysSerIysPhe 240
728 ATACATCTACAGCTGCCGCACTCCCTTATCATGATGATCAACCGCT 777
   : : : : : : : : : : : : : : : : : : : : : : : : : :
240 sPasnLysTyrGlyLysArgHisSerLeuIleAspGlyIleAsnArgAla 256
778 TCCGATGTCATATGGCGGCAAGACACTCTGTCATGGGTTACGGCGA 827
   : : : : : : : : : : : : : : : : : : : : : : : : : :
257 ThrAspValLeuIleGlyLysThrAlaValAlaPheGlyTyrGlyAs 273
828 TCTGCGAGAGGCTGCGTCATCCCTCCGTCGCAAGGCGCTGCTA 877
   : : : : : : : : : : : : : : : : : : : : : : : : : :
273 pValGlyLys...AlaAlaArgSerLeuArgGlyGlnGlyAlaTyrVal 289
878 TCATCAGAACTGCAACCAATCTGCGCTCCAGGCTGCCATGGAAGC 927
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

289 lValAlaThrGlnLeuAspProIleCysAlaLeuGlnAlaIleMetAspGly 305
928 TACGAGTCCCGCGCATGAGAGTCGTCACAGAGTGCATATCTTCGT 977
   : : : : : : : : : : : : : : : : : : : : : : : : : :
306 TyrGlnValAlaThrLeuAspGlnValAlaAspLysAlaSerIlePhe 322
978 TACATGCAACAGAACTGCCATATCATCTCTGTCATGATGCGCAGA 1027
   : : : : : : : : : : : : : : : : : : : : : : : : : :
322 eThrThrThrGlnLysAsnLysAspIleIleMetAlaIleAspMetAla 339
1028 TGAAGATTAAGCTATTCGCGTACATGCGCCACTTCGATTAAGAAAT 1077
   : : : : : : : : : : : : : : : : : : : : : : : : : :
339 eLysHisGlnAlaIleValGlyAsnIleGlyHisPheAsnGlnIle 355
1078 GATACAGATGGCTCATGAAATATCCAGGATGACAGACATCCCAATCA 1127
   : : : : : : : : : : : : : : : : : : : : : : : : : :
356 AspMetAlaGlyLeuAlaLysValProGlyIleValLysAspGluIle 372
1128 GCCAGAAATACGACATGTGGAAATCCAGATGGCCAGCATTCCTTC 1177
   : : : : : : : : : : : : : : : : : : : : : : : : : :
372 sProGlnValHisThrTyrThrPheProAspGlyLysIleIleVal 389
1178 TTGCTGAGGCGCGCTTCATACCTTGCCTGCGCTACAGTCAACATCT 1227
   : : : : : : : : : : : : : : : : : : : : : : : : : :
389 eUserGlnGlyArgLeuLeuAsnLeuGlnLysAlaThrGlnHisProSer 405
1228 TTGCTTATGTCATGTCAATTCACAAACAGACACTGCTCAGCTGAC 1277
   : : : : : : : : : : : : : : : : : : : : : : : : : :
406 PheValMetSerAsnSerPheAlaAspGlnThrLeuAlaGlnIleGlu 422
1278 CTACGAAAGAGAGAAATCTCGAGAGAAAGTTCACACTTCGCAAGC 1327
   : : : : : : : : : : : : : : : : : : : : : : : : : :
422 uPheThrLysProGlyAspTyrProThrGlyValTyrThrLeuProLys 439
1328 ATTCGATGAGAAAGTCCGCTCCGCTCAGCTCGATCTCGATGTCAC 1377
   : : : : : : : : : : : : : : : : : : : : : : : : : :
439 lIleAspGlnLysValAlaArgLysHisLeuAspAlaLeuGlnTyrG 455
1378 CTACAAAGCTTACACAGAGGCTGACTCATGATCAAGTTCGATGTA 1427
   : : : : : : : : : : : : : : : : : : : : : : : : : :
456 LeuThrThrLeuArgProGlnGlnAlaSerTyrIleGlyValAspVal 472
1428 GCGTCTTACAAAGTGTGATGCTTACCGTTAT 1458
   : : : : : : : : : : : : : : : : : : : : : : : : : :
472 pGlyProTyrLysSerAspHisTyrArgTyr 482

seq_name: sp_bacteriap:Q9CCJ4
seq_documentation_block:
ID Q9CCJ4 PRELIMINARY; PRT; 492 AA.
AC Q9CCJ4:
DT 01-JUN-2001 (TREMblrel. 17, created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE PUTATIVE S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE.
GN SAH OR M0771.
OS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Bacteria; Firmicutes; Corynebacterineae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elismeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Hotroyd S., Hornsby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.,

```


"Massive gene decay in the leprosy bacillus." ;
1988 1007 1011 (2001)

KL Mature 409:100/-1011(2001).
DR1 EMBL; AL583919; CAC30280.1; -.
DR HSSP; P10760; IDAF.
DR Leproma; M0771; -.
DR InterPro; IPR000043; Ado_hcysase.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; ADOHCYASR_1; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 492 AA; 53386 MW; DBB3DA54CF87EBB5 CRC64;

	alignment_scores:	
Quality:	1537.50	Length: 477
Ratio:	3.932	Gaps: 3
Percent Similarity:	81.971	Percent Identity: 62.266

```
alignment_block:
US-09-759-990-1 x Q9CCJ4 .
```

Align seg 1/1 to: Q9CCJ4 from: 1 to: 492

34 GAGTACAGAAATGGCCAAATCAACATCGATGTTCTCCGGCCCAAGAAGACT 83
17 ASPRPhelysValAlaAspIleuSerLeuAlaAsnIleuValTrpLysIle 33
84 TACCCCTGGTCGAGAAAGAAATGCCAGGCTTATGCTCTTCGTGAGCGTT 133
33 uAspIleuAlaIleuTrpIleuMetProGluLeuMetSerIleuAlaGlnIle 50
134 ATTCCGCTGTAAACCAATTGAAGGCTGCAGAAATCTGTGGTTCCCTCCAC 183
50 yTrIaGluValGlnProIleuLysGluAlaArgIleSerIleuSerIleuAla 66
184 ATGACAGTCCAGACAGCGCGTCTGATCCAGACATCCACAGCTCTTGCTG 233
67 MetThrValGlnTrpAlaValIleuIleuLeuTrpIleuThrAlaLeuVal 83
234 TGAATGCATAGGGCTCTCCACAACTCTCTCAACAAGATATCAAGCCG 283
83 aGluValArgTrpAlaSerGlyAsnIlePheSerTrpGlnAsnIleAla 100
284 CTGCTGCTATCGTTGTGCGCCCAACAGGCACACACAGAAAGCCAGCCGT 333
100 IaaIaIaValaIaValaIeIyProGlyTrpIleProGluIleuTrpLysGlu 116
334 ATCCCAATCTCCGCTGGAGAGGCCCAACACATCCAGATACAGTCAGAGAA 383
117 ValTrpValIleAlaTrpLysGluIleuTrpLeuGluIleuTrpTrpAl 133
384 CACATACCGCGCTGCATGATGGCCAGATGGTCAAGGCCCA...CAGCAGG 430
133 aAlaGluGlnMetLeuThrTrpProAspTrpAspLysProValAsnMetI 150
431 TTGTGCAATGATGGTGGTATGCTAACACTCTCATCTCCAAAGGCTTCGAA 488
150 IeIeuAspAspGlyLysIAspAlaTrpMetLeuValIleuArgIleValGln 166
481 TTTCGACAGCGCGGCTGTTCACAGGCCCAACAGAAAGTGCACAACTCGA 538
167 TyTrIuLysAlaGluValaIValaProTrpAlaGluValaAspAspSerAlaG 183
531 ATACCGGTGCGCTTCTTCTACATCAAGCAGGCTTCAACCAAGACAGA 586
183 uTrpLysValaIleuLeuAsnIleuValArgLysArgPheGluTrpAspLysG 200
581 ACCACGTGGCACACATGCTGCTGCCGCGCATCAAGGTGTTTCCAAAGAGACA 633
200 IuLyStrpThrLysIleAlaLysSerValaLysGluValaThrGluGluThr 216
631 ACAACAGGTGCCACCGGCTGTACACAGTCGACAGAAAGAGGCGCAACTCT 680
217 ThrTrpGluValIleuArgLeuTrpGlnPheAlaIaIaIaGlyAspLeuAl 233

```

681 CTCGCCACCACTCAACGCTCAACAGCGCTCTTCAAGATCCAACTGTGATA 730
233 aphetProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspa 250
731 ACATGTACGGGCGCGCACTCCCTTATGATGGTATTCAAACGCTGCTCC 760
250 snLysTyrGlyThrArgHisSerLeuIleAspLylLeuAsnArgLylThr 266
781 GATGTCAATGATCGCGCGCAAGACGCTCTTCATGGGTACCGCGCATGT 830
267 AspSerLeuIleGlyLysAsnValIleuIleCysGlyTyrGlyAspVa 283
831 CGGCAAGGGCTCGGCTCAATCCCTCGGCGCCAAAGCGCGCTGGTTATCA 880
283 LglLysSclCysAlaGluAlaIalAlaLysGlyGlnGlyAlaArgValThrI 300
881 TCACGAACCTCACCCCAATCTGGCTCTCCAGGCTCGCATGAAAGGCTAC 930
300 IethrGluIleAspProIleAsnAlaLeuGlnAlaLeuMetGlnGlyPhe 316
931 CAGGTCCGCGCATCGAGAAGTCGTCAGAGATGTCATATCTTTCGTTAC 980
317 AspValLysArgValGluAspValIleAlaAspSerAspIleValThr 333
981 ATGCACAGAACTCGGATGATCATCTCTGTGACATGATGCGCCAGATGA 1030
333 TalathrGlyAsnLysAspIleIleLeuLeuGlnHisMetLysAlaMetL 350
1031 AGGATTAAGGCTATTGTGCGTAACATCGCGCCACTTCGATTAAGAAATTGAT 1080
350 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 366
1081 ACAGATGGCTCATGAATAATCCAGGATCAACAGACATCCCAATCAAGCC 1130
367 MetAlaAlaLeuGlnArg...SerGlyAlaThrArgLeuAsnIleLysP 382
1131 AGAATACGACATGGGGAATTCGCCAGAT...GGCGACGCTATCCTCCTTC 1177
382 odIlnValAspLeuThrPheGlyAspSerGlyLysSerIleLeuValL 399
1178 TTGCTGAGGGCGCGCTTCTTAACCTTGGCTCGGCTACAGGTCACCCATCT 1227
399 euserGluLysArgLeuLeuAsnLeuGlyAsnAlaThrGlnHisProSer 415
1228 TTGCTATGTCAATGTCATTCACAAACACACACATCCTCGTCAAGCGAAGCT 1277
416 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 432
1278 CTACGAAAGAGAGAAATCTCGGACAGAAAGGTTTACACATCTCCGAAGC 1327
432 urThrLysAsnAspAspLysArgAspAsnGluValTyrArgLeuProLysH 449
1328 ATCTCGATGAAGAATGCGCTCGGCTCCACCTCGGATCTCTCGATGTCCAC 1377
449 IIsleuAspGluLysValAlaIaArgValHisValGluAlaGlnGlyGlyGln 465
1378 CTTACAAAGGTTTACACAGAGAGGCGGATCATCAATCAACGTTCCAGTTGA 1427
466 LeuThrLysLeuThrLysAspGlnAlaGluTyrIleuGlyValAspValAs 482
1428 GGGTCTTCAACATGTCGATCTTACCGTAT 1458
482 pGlyProPheLysProAspHisTyrArgTyr 492
seq_name: sp_plant:O9LK36
seq_documentation_block:
ID O9LK36 PRELIMINARY; PRT; 485 AA.
AC O9LK36
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
HYDROLASE) (ADHOCYASE) (S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE).
GN MY9.16
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA MEDLINE=20363099; PubMed=10907853;
RX Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl.
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Chouk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinzaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
CC + L-HOMOCYSTEINE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
DR EMBL: AP000377; BAB01858.1; -;
DR HSSP: P10760; 1B3R.
DR InterPro: IPR000043; Ado_hcyase.
DR InterPro: IPR000860; porphobil_deam.
DR Pfam: PF00670; AdhCysase.1.
DR PROSITE: PS00738; ADHOCYASE_1; 1.
DR PROSITE: PS00739; ADHOCYASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 485 AA; 53159 MW; 876079E94DC3732 CRC64;

alignment_scores:
Quality: 1512.00 Length: 487
Ratio: 3.887 Gaps: 4
Percent Similarity: 79.877 Percent Identity: 61.396

alignment_block:
us-09-759-990-1 x Q9LK36 ..

Align seg 1/1 to: Q9LK36 from: 1 to: 485

```

19 ACTGGTGCATTCGATGACAGATGCGACATCAACATCCATCTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThirserseclYargsluYrlyValysAspmetSerGlnAlasph 24
69 GCGCGTAAGAACTTACCTTGTGTGAAGAAATGCCAGGCTTTATGG 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
24 eglYrleuglullegluldeuAlaIvalGluMetProglYleuValS 41
119 TTCTTGATGAGGTATTCGCTTCTTAAGCATGGAAGGCTGTCAATC 168
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

41 erCysValThrGluPheGlyProSerGlnProLeuYsglyAlaArgIle 57
169 TTCTGTTCCCTCCACATGACAGTCGACAGCGCGCTCCATCGAGACT 218
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
219 CACAGCTTTGGTGTGATGATGAGATGGGCTTCCTGCACATCTTCTTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
74 utThrAlaLeuGlyAlaGluValAlaArgTrpCysSerCysAsnIlePheSer 91
269 CACAAGATACAGCGCGCTGCTATGCTGTTCGCGCCCAACAGCAGACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
91 hrGlnAlaSerHisAlaAlaAlaAlaAlaAlaArgAspSerAlaAla 105
319 GAGAAGCCAGCGCGGTATCCAGCTTTCGCTGAGGAGGCGCAACATGCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 .....ValPheAlaIleTrpLysGlyGluThrLeuG1 115
369 AGAATFACTGGAGAAACATATACCGCGCTCTACATGGCCAGATGTCAMG 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
115 nGluTrpTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132
419 GCCCAGCAGAGCTTGTGATGATGATGGTGGATGCTACATCTCATCTCC 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 lYProAspLeuIleValAlaAspAspGlyAlaAspAlaThrLeuLeuHis 148
469 AAGGCTTCGAA.....TTCCAACAGCGCGGTGCTGTTCCAGA 506
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
149 GluGlyValLysAlaGluGluIlePheAlaLysAsnGlyThrPheProAs 165
507 GCCACAGAGAGCTGACAACTCGAATATACCGTGGCTTGTTCCTACACTCA 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
165 pProThrSerThrAspAsnProGluPheGluIleValLeuSerIleIle 182
557 AGCAGCTCTTCAACCAAGACAAGAACCACTGGGACACAGTTCGTCGCCGC 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182 YsAspGlyLeuGlnIleValAspProLysLysThrHisLysMetLysGluArg 198
607 ATGAACGGTGTTCCTCGAAGAGACAACAACAGGTGTCACCGCTTCACCA 656
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
199 LeuValGlyLysAlaSerGlnGluThrThrThrGlyValLysArgLeuYrgI 215
657 GCTCGAAGAGAGGCGCAACTCTCTTCCACGCATCAACGTCAACAGCAG 706
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
215 nMetGlnGluThrGlyAlaLeuPheProAlaIleAsnValAsnAspS 232
707 CTGTACAAAGTCCAGTTCATATACATTCAGGCGCGCCATCTCCCTT 756
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 erValThrLysSerLysPheAspAsnLeuYrgLysArgHisSerLeu 248
757 ATCGATGATATCAACCGTTCGATGATGATGATGATGATGATGATGATG 806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 ProAspGlyLeuMetArgAlaIleThrAspValMetIleAlaGlyLysValAl 265
807 TCTCGTCATGGGTACGCGCATGTCGCGCAAGGCTGCGCTCAATCCCTCC 856
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 aValIleCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaIleMetL 282
857 GTGGCCAAGCGCGCTGCTATCATCAACAGAACTGACCAATCTCCGCT 906
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
282 YsThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCGAGGCTCCCATGGAAGCTTACAGAGTCCGCCCATCGAAGAACTGCT 956
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315
957 CAAGATGTCATATCTGTTACATGACAGGAATCGCATATCATCT 1006
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
315 lSerGlnAlaAspLlePheCysThrThrThrGlyLysnLysAspLleIleW 332
1007 CTGTGACATGATGCCCATGGAAGTGAAGTGAAGGCTATGTCGCTGATCATC 1056
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 etValAspHisMetArgLysMetLysAsnAsnAlaIleValLysAsnIle 348

```

```

1057 GGCACCTTGATACGAATGATACAGATGGCTCATGAATACCCAGG 1106
|||||
349 GYHISPhEAsPrAsnglulIeAsPmeLleuGlYleuThrTYrProGl 365
1107 CATCAAGCATCCCAATCAAGCAATAGACATGTGGGATTCACG 1156
|||||
365 YVALLYSATgllEhrlleYsProGlnThrSPArgrTYrValPheProA 382
1157 AAGGCACGCT...ATCCCTCTCTTCTGATGAGGCGCCCTTCTTAACCTT 1203
|||
382 SPThrAsnSerGlYlelleValleuIaGlulYArgLeuMetAsnleu 398
1204 GCGTGGCTACAGATCACCATCTTTCGTATGATCATATGCATACAAA 1253
|||||
399 GLYcysAlaThrglYhISProSerPheValMetSerYsSerPheThAs 415
1254 CGAGACATCGCTCAGCTCAGCTCTACGAAGAAGA.....GGAAATC 1297
|||||
415 nGlnValIleAlaGlInleuGlYleuThrPAsnglulYsSerSerGlYlYst 432
1298 TCGAAGAAGGTTTACACACTTCGAGACATTCGATGAAAGACTCGCT 1347
|||||
432 YrGlulYsYsValTYrValleuProYhISleuAsPglulYsValAla 448
1348 CGGCTCCACCTGGATCTCGATGTCACCTTACAAAGCTTACACAGA 1397
|||||
449 AlAleuHISleuGlYsleuGlYAlaYrleuThrYsleuThrYsAs 465
1398 GCAGGCTGACTACATCAAGCTTCAGGTGAGGCTCTTACAGTCTGATG 1447
|||||
465 rGlnSerAsPtyrValSerIleProValGlulYrProTYrYsProValH 482
1448 CTTCACGCTTAT 1458
|||||
482 IsTYrArTYr 485
seq_name: sp_plant:0944K5
seq_documentation_block:
ID_0944K5 PRELIMINARY: PRT: 485 AA.
AC 0944K5:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE AT3G23810/MYB9.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.W., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Ondera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southgate A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RT Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF428329; AAL16259.1;
SQ SEQUENCE 485 AA; 53129 MW; 8AA4719B3FBFAF6 CRC64;

```

alignment_scores:

```

Quality: 1512.00 Length: 487
Ratio: 3.887 Gaps: 4
Percent Similarity: 79.877 Percent Identity: 61.396

```

alignment_block:

```

US-09-759-990-1 x 0944K5 ..
Align seg 1/1 to: 0944K5 from: 1 to: 485
19 ACTGGTCTCCATTCGATGATACGAATTCGCAATCAATCACTCCATCTCT 68
|||||
8 ThrSerSerGlYArgGlulYrYsValYsAsPmeSerGlAlaAsPh 24
69 CGGCGCTAAGAACTTACCTCTCGAGAGAAGAAATTCGAGCTTATG 118
|||||
24 eGlYArgleuIleleGlulYleuAlaGlulYleuMetProGlYleuValS 41
119 TTCTTCGTAGCGCTTATTCGCTTCTTAAGCCATTAAGGCTGTCAATC 168
|||||
41 eYcysValAlaGlulPheGlYProSerGlProleuYsGlYAlaArgIle 57
169 TCTGGTCCCTCCACATGATCCAGACAGCCGCTCTATGAGACACT 218
|||||
58 ThrGlYSerleuHISMetThrlleGlInThrlAlaValleuIleGlulThrl 74
219 CACAGCTCTTGCTGATGTCAGATGGGCTCTGCAACATCTTCTTA 268
|||||
74 urThAlaIleuGlYAlaGlulValArgTYrPcysSerYsAsnIlePheSer 91
269 CACAAGATACAGCCGCTGCTATGCTGTGCGCCCAACAGGACACCA 318
|||||
91 hrglAsPrlAlaAlaAlaIleAlaIleAlaArgAsPserAlaAla..... 105
319 GAGAAGCCAGCCGATGCCAGTCTGCTCGTGAAGGGGAACACTCC 368
|||||
106 .....ValPheAlaTrpYsGlYleuThrleuGl 115
369 AGAATACTGGAGAGACATACACATACCGGCTCTCAATGCGCAGATGCAAG 418
|||||
115 nGlulYrTrpTYrPcysThrlArgAlaIleuAsPTrglYrProGlYlYg 132
419 GCCCAGACAGGTTTCGATGATGTTGGTGAAGCTCACTCCATCTCC 468
|||||
132 YrProAsPleuIleValAsPAsPrglYAsPAlaThrlleuIleHIS 148
469 AAGGCTTCGAA.....TTCGAACAGCCGCTCTGCTTCAGCA 506
|||||
149 GlulYValYsAlaGlulGlulPheAlaIleYsAsnglYlThrPheProAs 165
507 GCCAACAAGCTGACAACCTCGAATACCGCTGCTTGTTCGATCACTCA 556
|||||
165 ProThrSerThrAsPAsnProGlulPheGlInIleValleuSerIlelel 182
557 AGCAGGTTTCAACCAAGACAGAACCACTGACACACAGTGTGCGCGC 606
|||||
182 YsAsPglYleuGlInValAsPProYsYsYrThISYsMeLYsGlulArg 198
607 ATGAACGCTGTTCGAGAGAGACAACACAGGTGTCACCGCTTACCA 656
|||||
199 leuValGlYAlaSerGlulGlulThrlThrlnglYValYsArgleuYrGl 215
657 GCTCGAAGAAGGCGCAACTCTCTTCCAGCCATCAACGTCAACAGCAG 706
|||||
215 nMetlInglulThrlglYAlaIleuIleuPheProAlaIleAsValAsnSPS 232
707 CTGTACAAAGCTCAAGTTCGATTAACATCAAGCTGCGCCACTCCCTT 756
|||||
232 eValThrlYsSerYsPheAsPAsnleuYrGlYsAlaHISSerleu 248
757 ATCGATGGTATCAACCGTCTTCGATGTCATGATGCGCGCAAGACAGC 806
|||||
249 ProAsPglYleuMetArgAlaThrAsPValMetIleAlaGlYsValAl 265
807 TCTGCTCATGGCTTACGGCGGATGTCGCAAGGCTGCGCTCAATCCCTCC 856
|||||
265 aValIleCysGlYTYrGlYAsPValGlYsGlYsAlaAlaIleMetL 282
857 GTGGCAAGGCGCTGCTTATCATCAAGAACTCGACCAACTCTGGCT 906

```

```

287 yThralaGlyAlaValIleValIleThrGluIleAspProIleGlySala 298
290 CTCAGAGGTCGCATGAGAGGCTACAGGTCGCCGCGCATGAGGAAGTCGT 956
291 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
292 LeuGlnAlaLeuMetGluGluGluValLeuThrLeuGluAspValVa 315
293 CAAGATGTCGATATCTCTGTTACATCCACAGAAATGCGATATCATCT 1006
294 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
295 1SerGlnAlaAspIlePheCysThrThrThrYAsnIleAspIleIleM 332
296 CTGTGACATGATGCGCCAGATGAAGATAGGATGATGTCGTTACATC 1056
297 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
298 etValAspHisMetArgGlyMetGlySAsnAlaIleValCysAsnIle 348
299 GGCACATTCGATACGAAATGATATACAGATGAGTCATGAAATACCCAG 1106
300 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 GlyHisPheAspAsnGluIleAspMetLeuGluGluThrTyrProI 365
302 CATCAAGACATCCCAATCAAGCCAGAAATACGACATGTCGGAATCCAG 1156
303 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
304 yAllyAspArgIleThrIleLeuProGlnThrAspArgTyrValPhePro 382
305 ATGGCCACGCT...ATCCCTCTCTGTGTCAGGCGGCTCTCTTAACCTT 1203
306 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
307 sPThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
308 GCTGCGCTACAGATCACCACATCTTCTGTTATGTCATGATCATTCACAA 1253
309 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
310 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
311 CGACACATCTGCTGACGTCGACCTCTTACAGAAAGAGA.....GGAATC 1297
312 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
313 nGlnValIleAlaGlnLeuGluLeuThrPAsnGluLeuSerSerGlyLys 432
314 TCGAGAGAAGGTTTACACACTTCGAGATCCGATCGATGAGAGAGTCGCT 1347
315 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
316 yGluGlyLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
317 CGCTCCACCTCGATCTCTGATGTCACCTTACAAAGCTTACACAGAA 1397
318 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
319 AlaLeuHisLeuGluLysLeuGluAlaArgLeuThrLysLeuLysAs 465
320 CGAGGTCGATACATCAAGCTTCAGTTGAGAGGCTCTTACAAAGTCGATG 1447
321 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
322 pGlnSerAspTyrValSerIleProValGluGlyProTyrLysProVal 482
323 CTTCACGCTTAT 1458
324 ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
325 1sTyrArgTyr 485
326 seq_name: sp_plant:Q9SP37
327 seq_documentation block:
328 ID Q9SP37 PRELIMINARY; PRT; 485 AA.
329 Q9SP37;
330 01-MAY-2000 (TREMBLrel. 13, Created)
331 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
332 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
333 ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
334 DEHYDROLASE) (ADHCYASR).
335 SHH.
336 Lupinus luteus (yellow lupine).
337 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
338 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
339 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
340 NCBI_TaxID=3873;
341 [1]
342 SEQUENCE FROM N.A.
343 TISSUE=ROOTS INFECTED WITH BRADYRHIZOBIUM LUPINI;
344 Janowski R., Podkowiński J., Kisiel A., Jaskolski M.;
345 "S-adenosyl-L-homocysteine cDNA sequence from Lupinus luteus.";
346 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

```

CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
CC + L-HOMOCYSTEINE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
DR EMBL: AF185635; AAD56048.1; -.
DR HSPSP; P10760; 1B3R.
DR InterPro; IPR000043; Ado_hcyase.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR000860; Porphobil_deam.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; ADHCYASE; 1.
DR PROSITE; PS00739; ADHCYASE_2; 1.
KW Hydroxylase; NAD; One-carbon metabolism.
SQ SEQUENCE 485 AA; 53326 MW; 42F8PA6C58A42B19 CRC64;

```

alignment_scores:

```

Quality: 1506.00 Length: 482
Ratio: 3.862 Gaps: 4
Percent Similarity: 80.913 Percent Identity: 61.618

```

alignment_block:

```

US-09-759-990-1 x Q9SP37

```

```

Align seg 1/1 to: Q9SP37 from: 1 to: 485

```

```

34 GAGTACAGAAATGCGCAGATCAACCTCCATGCTTCTCGCCGTAAGGAAT 83
||||| ||||||| ||||||| ||||||| ||||||| |||||||
13 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuI 29
84 TACCTTCTGTCAGAGGAATGCCAGGTCCTTATGCTTCTTCTGACCGTT 133
||||| ||||||| ||||||| ||||||| ||||||| |||||||
29 eGluLeuAlaGluValAlaGluMetProGlyLeuMetAlaSerArgSerGlu 46
134 ATCCGCTTGTAAAGCATGTAAGGTCATGAGATGTCGTCGTCGTCGTC 183
||||| ||||||| ||||||| ||||||| ||||||| |||||||
46 hGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 62
184 ATGACAGTCCAGACAGCGCTCTCATCGAGACATCAGCTCTTGCTG 233
||||| ||||||| ||||||| ||||||| ||||||| |||||||
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlu 79
234 TGATGTCAGATGGGCTTCTGCAACATCTTCTTACACAGATPACAGCC 283
||||| ||||||| ||||||| ||||||| ||||||| |||||||
79 aGluValArgTyrPcysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATCGTGTGCGCCCAACAGGACACACAGAGAGGACGCGGT 333
||||| ||||||| ||||||| ||||||| ||||||| |||||||
96 IaLaLaIleAlaIleArgAspSerAlaAla..... 105
334 ATCCAGATCTTCTGCTGGAAGGCGGAACACTCCAGATATAGTGGAGAA 383
||||| ||||||| ||||||| ||||||| ||||||| |||||||
106 .....ValPheAlaIlePyrGlyGluThrLeuGlnGluTyrTyrPrcy 120
384 CACATACCGCGCTCTCAATGCGCAGATGTCAGAGGCGCACAGCAGTTG 433
||||| ||||||| ||||||| ||||||| ||||||| |||||||
120 sThrGluArgAlaLeuAspTyrGlyProGlyGlyGlyProAspLeuIle 137
434 TCGATGATGATGGATGATACATCTCCATCTCCAGAGGCTTGGAA... 480
||||| ||||||| ||||||| ||||||| ||||||| |||||||
137 AlaAspAspGlyGlyAspThrThrLeuLeuIleHisGluGlyValLysAla 153
481 .....TTGAAACAGCGGCTGCTGTCAGAGGACCAAGCAACCTGA 521
||||| ||||||| ||||||| ||||||| ||||||| |||||||
154 GluGluIleTyrGluLysSerGlyGlnPheProAspProAspSerThrAs 170
522 CAACCTCGAATACCGGCTGCTGTTCTTGTACACTCAAGCAGGTCCTCAACC 571

```


13 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuGluI 29
84 TACCGTTGGTGAAGAAGAAATGCCAGGCTTATGTTCTTCTGAGAGCTT 133
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaSerArgAlaGluP 46
134 ATTCGGCTTCTAAGCCATTGAAGGGTGTCAAAATCTGTGTTCCCTCCAC 183
46 heGlyProSerGlnProValLysGlyAlaLysIleThrCysSerLeuHis 62
184 ATGACAGTCCAGACAGCCGCTGCATGAGACACTCAAGCTCTTGAGTC 233
63 MetThrIleGlnThrAlaPheLeuIleGluThrLeuThrAlaLeuGlyAl 79
234 TGATGTCAGATGGCTTCCCTGCAACATCTTCTTACACAGATACAGCCG 283
79 agiValAlaGlyTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTGCTATGTTGTGCGCCCAACAGAGCACACAGAGAACCCAGCCGT 333
96 laIAlaIleAlaIleAlaArgAspSerAlaAla..... 105
334 ATCCAGTCTTGGCTGGAAGGCGAAACACTCCAGAAATAGTGGAGAA 383
106 ValPheAlaTrpLysGlyGluThrLeuGlnGluTyrTrpLys 120
384 CACATACCGCGCTTCAACATGCGCAGATGTCAGAGCCACAGCAGTTG 433
120 sThrGluArgAlaLeuAspTrpGlyProGlyGlyLysProAspLeuIleV 137
434 TCGATGATGGTGGTGTGATGCTACAGCTCATCTCCAAAGGGCTTC..... 477
137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153
478 GAATTCGAACACCGCGCTGCTTCCAGAGCCACAGAGCTGA 521
154 GluGluLupheGlnLeuValLeuThrIleLeuLysGluSerLeuLysT 170
522 CAACCTGGAATACCGCTGCTTGTGTACACTCAAGCAGAGTCTTCAAC 571
170 pasValGluPheGlnLeuValLeuThrIleLeuLysGluSerLeuLysT 187
572 AAGCAAGAACCACTGGCACAAGTGTGTCGCGCATGAACGGTGTTC 621
187 hrAspProLeuAlaGlyThrLysMetLysGluArgLeuValGlyValSer 203
622 GAAGAGACAAACAAGGTGTCCACCGCTTACAGCTCGAAGAGAGGG 671
204 GluGluThrThrThrGlyValLysLysLeuTyrGluMetProAlaAsnG 220
672 CAACCTGCTTCCAGCATCAAGCTCAAGCAGCTGTATCAAACTCA 721
220 ySerLeuLeuPheLeuProIleAsnValAsnAspSerValThrLysSerL 237
722 AGTTGATTAACATCTACGGCTGCGCCACTCCCTTATCATGTGATTAAC 771
237 yspheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGTTCGATGTCATGATCGCGGCGGCAAGACAGCTCTCGTCAATGCTTA 821
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaLeuValAlaGlyT 270
822 CGGCGATGTCGCAAGGGCTGCGCTCAATCCCTCGTGGCCAGGCGCT 871
270 rGlyAspValGlyLysGlyAlaAlaMetLysGlnAlaIleAla 287
872 GCGTATATCATCAGAACTCGACCAATCTGGGCTCCAGGCTGCGCATG 921
287 rGValIleValThrGluIleAspProIleCysAlaLeuGlnAlaThrMet 303
922 GAAGGCTACAGGTCGCGCATCGAAGAGAGCTCTCAAGATGTGATAT 971
304 GluGlyLeuGlnValLeuPheLeuGluAspValValSerGluValAspI 320

972 CTTCGTTATATGACAGAAACTGCAATATCATCTCTGTGACATGATGG 1021
320 ePheValThrThrThrLysAsnLysAspIleIleMetValAspHisMetAl 337
1022 CCCAGATGAAGATMAAGCTATTGTGCGTAACATCGCCACTTCGATTAAC 1071
337 rGlyMetLysAsnAsnAlaIleValLysAsnIleGlyHisPheAspAsn 353
1072 GAAATTTATACAGATGGCGCTCATGAAATCCAGCAGATCAACACATCCC 1121
354 GluIleAspMetHisGlyLeuGluThrPheProGlyValLysArgLeth 370
1122 AATCAACCCAGAAATATGACATGTGGAAATTCACATGGCCACGCT..A 1168
370 rIleLysProGlnThrAspArgTrpValPheProAspThrAsnSerGlyI 387
1169 TCCTCTCTTCTGCTGAGGGCGCTTCTTAACCTTGGCGCGCTACAGGT 1218
387 leIleValLeuAlaGluGluValArgLeuMetLysLeuGlyCysAlaTrpGly 403
1219 CACCATCTTCTGTTATGTCATGTCAATGCATCAACAAACAGACATCGCTCA 1268
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420
1269 GCTGACCTCTACGAAAGAGA.....GGAATCTCGAGAGAGATT 1312
420 nLeuGluLeuThrAsnGluArgSerSerGlyLysTrpGluLysValT 437
1313 ACACACTTCGGAAGCATCTCGATGAAGAGTGCCTGCTCCACTCGGA 1362
437 yValLeuProLysHisLeuAspGluLysValAlaIleAlaLeuHisLeuGly 453
1363 TCTCTGATGTCCACCTTACAAAGCTTACACAGACAGAGCTGACTATCAT 1412
454 LysPheGlyAlaLysLeuThrLysLeuThrLysAspIleAlaAspTyrI 470
1413 CAACCTTCAGTTGAGGGCTTACAAAGCTGTGATGCTTACCGTTAT 1458
470 eTyrValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485
seq_name: sp.plant:Q42939
seq_documentation_block:
ID Q42939 PRELIMINARY: PRT: 450 AA.
AC Q42939;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ADENOSYLHOMOCYSTEINASE (EC 3.3.1.1) (S-ADENOSYL-L-HOMOCYSTEINE
DE HYDROLASE) (ADOCYASE).
OS Nicotiana sylvestris (wood tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid=4096;
RN [1]
RP SEQUENCE FROM N. A.
RA Mitsui S., Wakasugi T., Sugiyama M.;
RT "A cDNA encoding the 57 kDa subunit of a cytokinin-binding protein
RT complex from tobacco : the subunit has high homology to S-adenosyl-L-
RT homocysteine hydrolase.";
RL Plant Cell Physiol. 34:1089-1096(1993).
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
CC ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINE MAY PLAY A KEY ROLE IN THE
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O - ADENOSINE
CC + L-HOMOCYSTEINE.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.

OC xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C.
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.U.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.U.S., Ferreira V.C.A., Fetto J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
 RA Garier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kurama E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Martins E.M.F., Matsukuma C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peloto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubaho M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-159(2000).
 CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-
 ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
 CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-HOMOCYSTEINE + H(2)O = ADENOSINE
 CC + L-HOMOCYSTEINE.
 CC -1- COFACTOR: NAD (BY SIMILARITY).
 CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
 CC DR EMBL, AE003941; AAF83647.1; -.
 DR HSSP, P10760; 1B3R.
 DR InterPro: IPR000043; Ado_hcyase.
 DR Pfam: PF00670; Adohcyase; 1.
 DR PROSITE: PS00738; ADOHCYASE_1; 1.
 DR PROSITE: PS00739; ADOHCYASE_2; 1.
 KM Complete proteome; Hydrolase; NAD, One-carbon metabolism.
 SQ SEQUENCE 446 AA; 49005 MW; F3B2P0F6311B9E3A CRC64;

alignment_scores:
 Quality: 1406.00 Length: 457
 Ratio: 3.790 Gaps: 5
 Percent Similarity: 81.182 Percent Identity: 60.613

alignment_block:
 US-09-759-990-1 x 09PEJ1 ..

Align seg 1/1 to: 09PEJ1 from: 1 to: 446

103 ATGCGAGCTTATGTTCTGTCGAGCGTTATTCGCGTTCTTAAGCAAT 152
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1 MetProcllyeMetserlleaIgrglgylgrAlaserlyglnProle 17
 153 GAAGGCTGTCAGAAATCTGTCGTCCTCCACATGACAGTCACAGACCGG 202
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 17 uLysglYvalaIgrYalThrIglSerleuNIsmethrIleIglInThrIalav 34
 203 TCGTCATGAGAGACAGTCACAGCTCTGTGTCGTCGATGTCAGATGGGCTTC 252

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 34 allleuilegluThrleuLysaspIleIglYalaspValaIgrTPrAlaser 50
 253 TGCACATCTTCCTTCATACAGATACAGCGCGTCGTCCTATCTTCGTCGCG 302
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 51 CysasnIlePheSerThrIglaspHisAlaIalalalaleIvalatrr.. 66
 303 CCCAACAGGACACACAGAGAGAGCGCGGTATCCAGTTCGTCGTCGGA 352
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 67 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 353 AGGCGAAACATCCAGATATCTGGAGACACATACCGCGCTTCACA 402
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 75 YsgIgluThrleuIgluIgluIgluIgluIgluIgluIgluIgluIgluIglu 91
 403 TGG.....CCAGATGT.....CAAGCCCAACAGCGGTTCGATGA 440
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 92 PheThrleuSeraspIgluThrleuThrIgluIgluIgluIgluIgluIgluIglu 108
 441 TGTGTGTATGCTACATCTCTCATCTCCAGAGCGCTTCGATTCGAAACAG 490
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 108 polYglYaspIalThrleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 125
 491 CCGGTGCT..GTTCAGAGCCACAGAGACCTGACAACTCGAATACCGC 537
 :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 125 IYserThrItrYalaspIgluPro.....SeraspIleuIgluIgluIglu 139
 538 TGCCTTCCTTCATACAGGAGGCTTCACACAGAGAGAGAGAGAGAGAGAGAG 587
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 140 ValIleYsArgleuIleuYsArgIleAlaIleIgluIgrProIgluYTr 156
 GCACACAGTGTGTCGCGGATGACAGCGGTTCGAGAGACACACACAG 637
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 156 PThrArgValValaspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 173
 638 GTGTCCACCGCTTCACAGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 173 IYvalHisArgleuYrIglIleAlaIalThrIgluIgluIgluIgluIgluIglu 189
 688 GCCATCAACGTCACAGAGCGCTGTTCAAGTTCAGATTCAGATCACTCA 737
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 190 AlaIleasnValaspIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 206
 738 CGGTCGCGCCACTCCCTTATGATGATGATCAACCGTCTCCGATGTC 787
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 206 rglYsArgIgluSerIleuAlaspIleuIleuIleuIleuIleuIleuIleuIleu 223
 788 TGTATGGGCGGACAGAGCTGTCGATGATGATGATGATGATGATGATGATGATGAT 837
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 223 eLleuAlaIglYsLleuAlaIValCysglYrIglYaspValIglYs 239
 838 GCGTCGCTCAATCCCTCCGTCGCGACAGCGCTCGCGTTCATCAACAG 887
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 240 GlYserAlaHisSerleuIgrAlaIgrAlaIgrAlaIgrAlaIgrAlaIgrAla 256
 888 ACTGCAACCAATCTGCGCTCCAGGCTGCGATGAGAGGCTACCAAGGCTCC 937
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 256 nIleaspIroIleCysAlaIleuIglAlaIametIgluIgluIgluIgluIgluIglu 273
 938 GCGCATGAGAGAGAGTGCATGAGATGATGATGATGATGATGATGATGATGATGAT 987
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 273 hTrhIvalIgluSprIleuIgluIgluIgluIgluIgluIgluIgluIgluIglu 289
 988 GGAACATGAGATATCATCTGTCGTCGATGATGATGATGATGATGATGATGATGATGAT 1037
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 290 GlYasnIlyaspValIleuIgluIgluIgluIgluIgluIgluIgluIgluIglu 306
 1038 GCGTATGTCGTCATCAATCGGCTTCGATTAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1087
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 306 nValIleValCysasnIleIgluIleuIleuIleuIleuIleuIleuIleuIleuIleu 323
 1088 GCGTCATGAATATCCAGGACATCAAGCATGATCCCATGAGCCAGATATAC 1137
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```

:::|||||
225 eutrycilylsarvgiuserlevalaspralaleatrgatgltthrasp 241
784 GTCATGATGGCGCGCAAGACAGCTCTGATGGGTTACGGCGATGTCGG 833
|||||
242 VALMELLESERGLYSLVALALAVALLCYSGLTYRGLYSPALVALGL 258
834 CAAGGGCTCGGCTCAATCCCTCGTGGCCAGAGCGCTCGCTATCATCA 883
258 yslgysleralaleaserleuarginglylalaarvallevalt 275
884 CAGAACTCGACCCCAATCTGGCTCTCCAGGCTCCAGGAAGCGCTACAC 933
275 hrglvalasprrollecyasalaueuinalalameglslytyrglu 291
934 GTCGGCGCATCGAGAACTCTCAAGAGTCTGATATCTTGTTACATG 983
292 valglthrlleuasnsapvalalaleasprlyalalephevalthral 308
984 CACAGAACTCGCATATCATCTCTGTGACATGATGGCCAGATGAAGG 1033
308 atthrglyasnllysapvallethvalasprmetarglysmelysa 325
1034 ATAGAGCTATTGTCGTAACATCGGCACTTCGATACAGAAATTGATACA 1083
325 snasnaalalevalcyasnlleghlynsrphesergerluleglnile 341
1084 GATGGCCTCATGAATAACCCAGGCATCAAGCAATCCCAATCAAGCCAA 1133
342 Alaglyleuarganphe.....lystrasprglulellyserol 355
1134 ATAGACATGTTGGGAATCCAGATGGCCAGCTATCTCTCTTGTGTG 1183
355 nvalhshisvalglupheprobspilylsylsleuilevalleuserg 372
1184 AGGCGCCCTTTAACTTGCGTGGCTACAGGTACCCATCTTTCGTT 1233
372 lugalyleuvalasnlleuglyasnaalathrglynsrpheserph 388
1234 ATGCATGTCATTCACAAACCCAGACATCGCTACGCTGACCTTACGA 1283
389 Metserlaserphethrasnlethrasnleuaglnilegluleutprh 405
1284 AAGAGAGGAATCTCGAAGAGAGTTTACACTTCCGAGACATCTCG 1333
405 rasnllysalaustyrgluasnglnvalytrthleuprollysnsleua 422
1334 ATGAAGAAGTGGCTCCACCTCGGATCTCGATGTCACCTTACA 1383
422 spglulysvalalalepheuhsleuglulysleuclalalalsleut 438
1384 AAGCTTACACAGAGCAGCTGATCATCACTTCCAGTTGAGGCTCC 1433
439 Thrleuaglylsasprlnalalspryrlleghlyvalproglulaglypr 455
1434 TTACAACTCTGATGCTTACCGTTAT 1458
455 orhelysprobasprhlslyrarglyr 463

seq.name: sp_bacteriap:Q92TCL
seq_documentation_block:
ID Q92TCL PRELIMINARY; PRT; 466 AA.
AC Q92TCL;
DT 01-DEC-2001 (TREMBLE rel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLE rel. 19, last annotation update)
DE PROBABLE ADENOSYLHOMOCYSTEINASE PROTEIN (EC 3.3.1.1).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Gallibert F., Flanar T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowers L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjel M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Leteure V.,
RA Masny D., Palm C., Peck M.C., Thebault P., Vandenbol M., Purnelle B.,
RA Rampeger U., Surzycki R., Thebault P., Wells D.H., Wong K., Yen K.-C., Batut J.,
RA Vorholter F.J., Weidner S., Wells D.H., Wong K., Yen K.-C., Batut J.,
RT "The composite genome of the legume symbiont Sinorhizobium meliloti."
RL Science 293:668-672(2001).
DR EMBL: AL591782; CAC41426.1;
KW Hydrolyase; Complete proteome.
SQ SEQUENCE 466 AA; 5062 MW; 81D2BC7AF63FC79 CRC64;

```

```

alignment_scores:
Quality: 1355.00 Length: 475
Ratio: 3.547 Gaps: 3
Percent Similarity: 80.421 Percent Identity: 57.263

```

alignment_block:

US-09-759-990-1 x Q92TCL ..

Align seg 1/1 to: Q92TCL from: 1 to: 466

```

34 GAGTCAAGATTTGGCGACATCAACCTTCATGTTCTGGCCGTGAAGACT 83
6 Aspyrlllevalalaspilleglyleuvalasprhleglyrlysglu 22
84 TACCTTCGTCGAGAGGAATGCCAGCTTTATGTTCTTCTGTCAGCGTT 133
22 ealallealaglutthrglumeprroglyleuvalalacysarglulup 39
134 ATTCGCTCTTAAGCCATTAAGAGGTCTCGATCTCTGTTCCCTCCAC 183
39 heglylaserlysprleuulysglyalalarglilethrglyserleuhs 55
184 ATGACAGTCAGACAGCCGCTCATCGAGAGACTACAGCTCTTGCTG 233
56 Metthrlleghlnthralvalaleuileghlnthrlleuvalalaleu 72
234 TGATGTCAGATGGCTTCTCGAATCTTCTTACACAGATTCAGCGG 283
72 aglulvalargtrpalasercysanllepheserthrglasprhlsala 89
284 CTGCTCTTATCTGTTGCGCCCAACAGACACAGAGAACCCAGCGGT 333
89 lalallealealala.....sergly 96
334 ATCCAGCTCTCGCTGGAAGGGGAAACATCCAGAACTAGGAGAA 383
97 valprovalalalevalalysglulthrlleucluglyutrgluprth 113
113 rthrasprlysllephelntrpalaasprglylvalaserlnmetlle 130
434 TCGATGATGGTGGATGATGATGATGATGATGATGATGATGATGAT 483
130 eunspasprglylaspalathrmetrylleleucluglyalalargla 146
484 GAAGAGCGCGGTCTGTCAGAGCCAGAGAGAGTGCACACTGAGATA 533
147 Glualaglyluasnlle.....leuthrasnprogllyserglucldu 161
534 CCCTGCTCTTCTTACATCAAGCAGCTCTTCAACCAAGCAAGAAC 583
161 uglulleuphealaglnllelylsyargleuvalalalatrproglly 178

```

[illegible]

87 CCTTGTGAGAGAAATGCCAGGTCTATGCTTCTGTCAGCGTTAT 136
 23 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluMetLys 40
 137 CGGCTTTAGCCATGTGAAGGGTGTGAGAAATCTGGTCCCTCCATG 186
 40 eAlaSerLysProLeuLysGlyAlaArgGlyLeuGlyCysLeuHisMet 56
 187 ACAGTCCAGACAGCCGTCTCATCGAGACATCCAGCTCTTGGTGTGA 236
 57 ThrValGluThrAlaValLeuIleGluThrLeuValAlaLeuGlyAlaG 73
 237 TGTCAAGATGGCTTCTGCAACATCTCTACACAGATATACAGCGGTG 286
 73 uValArgTrpSerSerCysAsnIlePheSerThrGlnAspHisAlaAla 90
 287 CTGCTATGCTTGTGGGCCAACAGGACACAGCAGAAAGCCAGCGGTATC 336
 90 lAlaIleAla.....LysAlaGlyIle 97
 337 CCAGTCTTGCCCTGGAAAGGCGAAACATCCCAAGATCTGGGAGAACAC 386
 98 ProValPheAlaTrpLysGlyGluThrAspGluGlyLysLeuTrpCysI 114
 387 ATACCGCGCTCACATGCCACATGGTTCAGGCCACAGCAGAGTGTGCG 436
 114 eGluGlnThrLeuHisPheLysAspGlyPro..LeuAsnMetIleLeuA 130
 437 ATGATGGTGGTATGCTACCTCTCATCTCCAAAGGCTTCGAATTCGAA 486
 130 sPaspGlyGlyAspLeuThrAsnLeuIle..... 139
 487 ACAGCGGCTCTGTCCAGAGCCAAAGAACTGACAACTCGAATACCG 536
 139 139
 537 CTGCGTCTTGTCTACACTCAAGCAGGTCTTCAACCAAGACAAAGCACT 586
 139 139
 587 GGCACACA.....GTGCTGCGCGCATGACGGTGTTCGGA 624
 140 ..HisThrLysTyrProGlnLeuLeuSerGlyIleArgGlyLysSerGlu 155
 625 GAGACAACACAGGTGTCCACCGCTCTACACAGCTCGAAGAGAGGCA 674
 156 GluThrThrThrGlyValHisAsnLeuTyrLysMetMetSerAsnGlyT 172
 675 ACTGCTCTTCCAGCCATCAAGTTCACAGCAGCTGTACAAAGTCCAACT 724
 172 eLeuLysValProAlaIleAsnValAsnAspSerValThrLysSerLysP 189
 725 TCGATACATCTAGGCTGCGGCCCATCTTATGATGGTATCAACGCT 774
 189 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyLysArg 205
 775 GCTTCGATGATGATCGCGGCAAGACAGCTCTGTCATGGGTTACG 824
 206 AlaThrAspValMetIleAlaGlyLysValAlaValAlaIleGlyTyrG 222
 825 CGATGTGGCAAGGGCTGCGCTCAATCCCTCGTGGCCCAAGGCGCTCGG 874
 222 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArgV 239
 875 TTATCATCACAGAACTGCAACCAATCTGCGTCTCAGAGCTCCATGAA 924
 239 aIleIleIleThrGluIleAspProIleAsnAlaLeuGlnAlaIleMetGlu 255
 925 GGTACACAGGTCCGCGCATCGAGGAAGTGTCAAGGATGTCGATATCTT 974
 256 GlyTyrGluValThrThrMetAspGluAlaCysLysGlyGlyAsnIlePhe 272
 975 CGTTACATGCACAGGAACCTGCAATATCATCTCTGTGACATGATGGCC 1024

||||| ||||| ||||| |||||
 272 eValThrThrThrGlyCysValAspIleIleLeuGlyArgHisPheGluG 289
 1025 AGATGAAGATAGGCTATTTGTGCTAATCGGCGGCATCTGCATACGA 1074
 289 lMetLysAspAspAlaIleValCysAsnIleGlyHisPheAspValGlu 305
 1075 ATTGATACAGATGGCTCATGAATATCCAGGACATCAAGCAGATCCCAAT 1124
 306 lIleAspValLysTrpPheAsnGlu..AsnAlaValGluLysValAsnI 321
 1125 CAAGCCAGAAATPACGACATGTGGAAATTCACAGATGGCCAGCTATCTCC 1174
 321 eLysProGlnValAspArgTyrTrpLeuLysAsnGlyArgArgIleIle 338
 1175 TTCTGCTGAGGGCGCTTCTTAACCTTGGGTGGGTACAGTCCCA 1224
 338 eLeuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 354
 1225 TCTTTCGTTATGTCAATGTCAATTCACAAACACAGACTGCTACGCTCA 1274
 355 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleG 371
 1275 CCTTACGAAAGAGAGGAAATCTGACAGAAAGTTTACACTTCCCA 1324
 371 uLeuTrpThrHisProAspLysTyrProValGlyValHisPheLeuPro 388
 1325 AGCATCTCGATGAGAAAGTGGCTGCTCCCTGACCTGATCTCGATGTC 1374
 388 yLysLysLeuAspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 404
 1375 CACCTTACAAAGCTTACACAGAGAGCGCTGATCATCAACGTTCCAGT 1424
 405 LysLeuThrLysLeuThrGlyLysGlnAlaGlnTyrLeuGlyMetProI 421
 1425 TGAGGCTCTTACAGTCTGATGCTTAACGCTAT 1458
 421 eAsnGlyProPheLysProAspHisTyrArgTyr 432


```

63 MetThrIleGlnThrAlaValLeuIleGlnThrLeuThrAlaLeuGlyPr 79
224 TGATTCAGATGGGCTCTCTGCAACATCTCTACCAAGATACAGCG 283
      :::::::::::::::::::::::::::::::::::::::::::::::
79 GlnValAlaArgTrpCysSerCysAsnIlePheSerThrGlnAspHisAla 96
      :::::::::::::::::::::::::::::::::::::::::::::::
284 CTGCTGCTATCGTGTGCGGCCCAACAGGACACACAGAGAAAGCCGCGT 333
      :::::::::::::::::::::::::::::::::::::::::::::::
96 LeAlaLeIleAlaArgAspSerAlaSer..... 105
      :::::::::::::::::::::::::::::::::::::::::::::::
334 ATCCAGTCTTGCCCTGGAAAGGCGAAACACTCCCAATATCTGGAGAA 383
      :::::::::::::::::::::::::::::::::::::::::::::::
106 .....ValPheAlaIleTrpLysGlnThrLeuGlnGlnIleTrpTrpCys 120
      :::::::::::::::::::::::::::::::::::::::::::::::
384 CACATACCGCGCTCTACATGGCCAGATGGTCAAGGCCACAGCAGAGTTG 433
      :::::::::::::::::::::::::::::::::::::::::::::::
120 sThrGlnAlaLeuAspTrpGlyProGlyGlyProAspLeuIleV 137
      :::::::::::::::::::::::::::::::::::::::::::::::
434 TCGATGATGGTGTGATGCTACACTCTCATCTCCAAAGGCTTC..... 477
      :::::::::::::::::::::::::::::::::::::::::::::::
137 AlaAspAspGlyGlyAspThrThrLeuLeuIleHisGlnGlyValLysAla 153
      :::::::::::::::::::::::::::::::::::::::::::::::
478 .....GAATTGAAACAGCCGCTGCTCCAGAGCCAAAGCAAGAGCTGA 521
      :::::::::::::::::::::::::::::::::::::::::::::::
154 GlnGlnGlnTrpGlnTrpGlnTrpGlnTrpGlnTrpGlnTrpGlnTrp 170
      :::::::::::::::::::::::::::::::::::::::::::::::
522 CAACCTCGAATACCGCTGCTCTGCTACATCAACAGCAGCTCTCAACC 571
      :::::::::::::::::::::::::::::::::::::::::::::::
170 PAsnAlaGlnPheGlnIleValLeuThrIleIleArgAspGlyLeuLysV 187
      :::::::::::::::::::::::::::::::::::::::::::::::
572 MAGACAAGACCACTGGCACACAGTTGCTCCGCGCATGACGAGTGTTC 621
      :::::::::::::::::::::::::::::::::::::::::::::::
187 AlaAspProThrLysTrpArgLysMetLysAspArgIleValGlyValSer 203
      :::::::::::::::::::::::::::::::::::::::::::::::
622 GAAGAGACAAACACAGCTGTCCACCGCTCTACCAAGCTCGAAGAGAGG 671
      :::::::::::::::::::::::::::::::::::::::::::::::
204 GlnGlnTrpThrThrThrGlyValLysArgLeuTrpGlnMetGlnAlaAsp 220
      :::::::::::::::::::::::::::::::::::::::::::::::
672 CAACCTCTCTCTCCAGCCATCAACGTCACAGCAGCTGTTCACAAAGTCA 721
      :::::::::::::::::::::::::::::::::::::::::::::::
220 nSerLeuLeuPheProAlaIleAsnValAlaAspSerValThrLysSerL 237
      :::::::::::::::::::::::::::::::::::::::::::::::
722 AGTTCGATTAACATCTACGCTGCCCGCCACTCCCTTATCGATGATCAAC 771
      :::::::::::::::::::::::::::::::::::::::::::::::
237 yAspPheAspAsnLeuTrpGlyCysArgHisSerLeuProAspGlyLeuMet 253
      :::::::::::::::::::::::::::::::::::::::::::::::
772 CGGTGCTCGATGATCATGATCGGCGCAAGAGCTCTCGTATGGGTGA 821
      :::::::::::::::::::::::::::::::::::::::::::::::
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValAlaCysGlyTrp 270
      :::::::::::::::::::::::::::::::::::::::::::::::
822 CGGCGATGTGCGCAAGGCTGCGCTCAATCCCTCGTGGCCAAAGGCGCTC 871
      :::::::::::::::::::::::::::::::::::::::::::::::
270 rGlnAspValGlyGlnGlyCysAlaAlaLeuLysGlnAlaGlyAla 287
      :::::::::::::::::::::::::::::::::::::::::::::::
872 GCGTTATCATCAGAACTGACCCCATCTGCGCTCTCCAGGCTCCCATG 921
      :::::::::::::::::::::::::::::::::::::::::::::::
287 rValIleValIleThrGlnIleAspProIleCysAlaLeuGlnAlaLeuMet 303
      :::::::::::::::::::::::::::::::::::::::::::::::
922 GAAGGCTACAGAGTCCGCGCATGAGAGAGTCTCAAGATGTGATAT 971
      :::::::::::::::::::::::::::::::::::::::::::::::
304 GlnGlyLeuGlnValLeuThrLeuGlnAspValAlaSerGlnAlaAspI 320
      :::::::::::::::::::::::::::::::::::::::::::::::
972 CTTCGTTACATGACAGAAACTCGATATCATCTCTGTTCAGATGATGG 1021
      :::::::::::::::::::::::::::::::::::::::::::::::
320 ePheValIleThrThrThrLysAsnLysAspIleIleMetLeuAspHisMet 337
      :::::::::::::::::::::::::::::::::::::::::::::::
1022 CCCAGATGAAGGATAGGCTATTTCTCGTATCATCGGCGCATTCGATAC 1071
      :::::::::::::::::::::::::::::::::::::::::::::::
337 rGlySerMetLysAsnAsnAlaIleValLysAsnIleGlyHisPheAspAsn 353
      :::::::::::::::::::::::::::::::::::::::::::::::
1072 GAATATGATACAGATGGCTATCAATAATCCAGGACATCAAGACATCCC 1121
      :::::::::::::::::::::::::::::::::::::::::::::::
354 GlnIleAspMetLeuGlyLeuGlnThrTrpTrpGlyIleLysArgIleHis 370

```

```

1122 AATCAAGCCAGAAATACGACATGTGGAAATTCAGAGGCGCAC...GCTA 1168
      :::::::::::::::::::::::::::::::::::::::::::::::
370 rIleLysProGlnThrAspArgTrpValPheProGlnThrAsnThrGlyI 387
      :::::::::::::::::::::::::::::::::::::::::::::::
1169 TCCTTCCTTCCTGCTGAGGCGCCCTTCTTAACTTGGCTGGCTACAGT 1218
      :::::::::::::::::::::::::::::::::::::::::::::::
387 lIleValLeuAlaGlnGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
      :::::::::::::::::::::::::::::::::::::::::::::::
1219 CACCACTCTTCTGTTATGTCAATGTCAATTCACAAACAGACACTGCTCA 1268
      :::::::::::::::::::::::::::::::::::::::::::::::
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 420
      :::::::::::::::::::::::::::::::::::::::::::::::
1269 GCTGACCTCTACAAAAGAGA.....GGAATCTCGAAGAAAGGTTT 1312
      :::::::::::::::::::::::::::::::::::::::::::::::
420 nLeuGlnLeuTrpAsnGlnLysAlaSerGlyLysTrpGlyLysValT 437
      :::::::::::::::::::::::::::::::::::::::::::::::
1313 ACACACTTCGGAAGCATCTGATGAGAAAGAGCGCTCCACCTCGGA 1362
      :::::::::::::::::::::::::::::::::::::::::::::::
437 yValLeuProLysHisLeuAspGlnLysValAlaAlaLeuHisLeuGly 453
      :::::::::::::::::::::::::::::::::::::::::::::::
1363 TCTCTCGATGTCCACTTACAAAGCTTACACAGAGAGGCTGACTACAT 1412
      :::::::::::::::::::::::::::::::::::::::::::::::
454 LysLeuGlyAlaLysLeuThrLysLeuSerProSerGlnAlaAspTrpI 470
      :::::::::::::::::::::::::::::::::::::::::::::::
1413 CAACGTTCCAGTTGAGGCTCTTACAAAGTCTGATCTTACCGTTAT 1458
      :::::::::::::::::::::::::::::::::::::::::::::::
470 eSerValProIleGlnGlyProTrpLysProHisTrpArgTrp 485
      :::::::::::::::::::::::::::::::::::::::::::::::
seq_name: /cgn2_6/plotdata/2/laa/5b_comb.pep:us-08-896-005-5
seq_documentation_block:
; Sequence 5, Application US/08896005
; Patent No. 5854023
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,005
; FILING DATE: Filed Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0337 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

```

LIBRARY: GenBank
CLONE: 904132
US-08-896-005-5

alignment_scores:
Quality: 1278.00 Length: 478
Ratio: 3.662 Gaps: 5
Percent Similarity: 73.013 Percent Identity: 56.485

alignment_block:
US-09-759-990-1 x US-08-896-005-5

Align seg 1/1 to: US-08-896-005-5 from: 1 to: 432

```

37 TACAGATTCCGACATCAACCTTCAGTTCGCGCGTAAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 TATLVAALAAAPRIAGLYLEUALAATRPGLYARGLYSALAILEUAS 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
87 CCTTGTGAGAGAAATGCAGGCTTATGCTCTTGTGAGCGTTAT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 PILEAAGLVAASGLMETPROGLYLEUMETARGMETARGGLUMETYS 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 CGGCTTCTAAGCCATGTAAGGAGTGTGAGAACTCTGTTCCCTCAATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
40 ERLASERLYSPROLEULYGLYALARGYLEAGLYCYSEUHISET 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
187 AAGATTCAGACAGCCGCTCTCATGAGACATCAAGCTTGTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 TIRVALGLTIRVALAVALLEULIEGLTIRLEUVALALEUGLYALA 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 TGTCAAGATGGCTTCTGCAACATCTTCTACACAGATACACCGCTG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
73 VALARGTRPSEISERCYSASNILEPESERTIRGLASPHISALAIAA 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
287 CTGTATCTGTGTGGCCCAACAGGCAACACAGAAAGCCAGCCGATC 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90 IAAIAIILEAIA.....LysAlaGlyIle 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
337 CCAGTCTTGGCGTGAAGGGCGAAGCACTCCGCAATACCTGGGAACAC 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
98 PROVALIRPHEALATRPYSGLYLUTIRASPGLULYIRLEUTRPSYII 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 ATACGGGCTTCACATGGCCAGATGGTCAAGGCCACAGACAGGTTGCG 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
114 EGLULIRPHEULHISPHELYASPRILEURO...LEUASMETILEUEA 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
437 ATGATGGTGGTATGCTACATCTCTCATCTTCAAGGCTTCGAATTCGAA 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
130 SPASPGLYGLYASPRILEUTIRASNILEUIE..... 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
487 ACAGCGGCTGTGTCCAGAGCAACAGAAAGTGAACACTCGAATACCG 536
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 ..... 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
537 CTGCGTCTTGTCTACACTCAAGCAGGTCTTCAACAGACAGAACCACT 586
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 ..... 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
587 GGCACACA.....GTTGCTCCCGCATGAACGGTGTTCGAA 624
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
140 ..HISTIRLYTIRPROGLINLEUENISERGLYLEARGGLYILESERGL 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 GAGACACAAGAGGTGTCCACCGCTTCACAGCTGACAGAGAGGCGAA 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
156 GLUTIRHTRHTRGLYVALHISASNILEUTRYSKMETSETSNGLYII 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
675 ACTCTCTTCCAGCATCAACGTCAACGACGCTGTACAAATCCCAAGT 724
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 ELEUASVALPROALILEASNVALAISNAPSERVALTIRLYSERLYSP 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
725 TCGATACATCTAGGCTGGCGCCATCCCTATGATGATCAACCGT 774
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

189 HEASPRASNILEUTRGLYCYSARGGLUSERILEASPRGLYIILEYSLARG 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
775 GCTTCCGATGTGATGATCGCGGCAAGACAGCTCTGCTATGGGTACGG 824
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
206 ALATIRASPVALMETIRLEAIGLYLSVALAVALAVALAIGLYTYRGL 222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
825 CGATGTGGCAAGGGCTGGCTGCTCAATCCCTGGCGCAAGGCGCTGGC 874
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
222 YASPRALIGLYSGLYCYSLAIGLINALAUEATRGLYPHNEGLYALAARG 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
875 TTATCATCACAGAACTGACCCCATCTGCGTCTCCAGCTGCATGAA 924
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
239 ALILEIETIRGLIUEASPRILEASPRILEASPRILEASPRILEASPR 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
925 GGTATCCAGATCCCGCGCATTCAGAGAGTCTGTAAGATGTCATATCT 974
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
256 GLYTYRGLIUEALITIRHTRMETASPRGLIUALACYSLGILUYSNILEPH 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
975 CGTTACATGACAGGAAACTGCGATATCATCTGTTGACATGATGGCC 1024
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
272 EVALITIRHTRHTRGLYCYSVALASPRILEIILEUGLYARGHISPHEN 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1025 AGATGAGGATPAGGCTATGTCGGTACATGCGGCCCATCTGCATGAA 1074
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
289 IMETLYASPRASPRALILEVALCYASNILEGLYHISPHASPRVALGL 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1075 ATTGATACAGATGGCTCATGAAATACCCAGGATCAACACATCCCAAT 1124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 ILEASPRALYSTRPLEUASNILEU...ASNALAVAGLULYSVALASNII 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1125 CAAGCCAGAAATACGACATGTGGAAATGCCAATGAGCCAGCATCTCC 1174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
321 ELYSPROGLINVALASPARGYTYRTRPLEUYSNGLYALGATGIIIELE 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1175 TTCTTGTGAGAGGCGCTTCTTAACTTGTGCTGCGCTACAGTCCCA 1224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
338 EULEUALAGLUGLYARGLEUVALASNILEUGLYCYSLAMETGLYHISPR 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1225 TCTTTCGTTATGTCATGTCACAAACAGACAGACATCGGTCATGCA 1274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 SERPHEVALMETSERASNSERTHERASNGLINVAIMETALAGILEGI 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1275 CCTCTACGAAAGAGAGAAATCTCGAGAGAGGTTTACACATTCGCA 1324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
371 ULUTRPTIRHISPROASPRYSLYRPROVALIGLYVALHISPHENLEU 388
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1325 AGCATTCGATGAGAGAGTCCGCTGCGCTCCACCTGGAATCTGATGTC 1374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
388 YSLYSLEUASPRGLIUALAVALAIGLUALAHISLEUGLYLYSLEUAS 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1375 CACCTTACAAAGCTTACACAGAGAGGCTGACTACATCAACAGTCCAGT 1424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 LYSLEUTHIRLYSLEUTHIRGLIULYSGLINALAAGINLYRLEUGLY 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1425 TGAAGGCTCTTACAGTCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
421 EASNGLYPROPHELYSPROASPRHISTYRATGYT 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:US-08-896-005-4

seq_documentation_block:

Sequence 4, Application US/08896005

Patent No. 5854023

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purnvi

TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,005
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 178277
US-08-896-005-4

alignment_scores:
Quality: 1255.00 Length: 478
Ratio: 3.617 Gaps: 5
Percent Similarity: 72.594 Percent Identity: 55.439

alignment_block:
US-09-759-990-1 x US-08-896-005-4 ..

Align seg 1/1 to: US-08-896-005-4 from: 1 to: 432

```

37 TCACAGATTCGCGACATCCATCCATGTTCTGCGCGTAAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 TYRYSVALALASPLIEGLYLEUALALATRPOLYARGYSLALEUAS 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 CCTTGCTGAGAAATGCGAGTCTTATGTTCTTCTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 PILEALAGLUAISGLUMETPROGLYLEUMETARGMETARGLUATGYIS 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 CGCGTTCTAAGCATGAGGTTCTCAGAACTCTGTTCCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
40 ERLASERYSPROLEULYSGLYALATRGYLEALAGLYCYLEUHSMET 56
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ACAGTCACAGACGCGCTCATGAGACACTCAGAGCTTGTGNGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
57 THVALGLUHRALVALLEULIEGLUHRLEUVALTHRLUGLYALAGL 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 TGTCAAGTGGCTTCTGCAACATCTTCTACACAAATACAGCCGCTG 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 UVALGINTRPSESRERYASNIHPESERTHGLINSNHSIALALAA 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 CMCCTATCGTGTGCGCCCAACAGACACACAGAGAACCCGCGTATC 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 IALALALEAA.....LYSLALAGLYLE 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 CCAAGTTCGCGCTGGAAGGCGCAACTCCACAGATACGAGAAACAC 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
98 PROVALYRHLATRPYSLGLUHRASPLUGLUYTRYLEUTRPOYSIL 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 ATACCGCGCTCTCACATGCGCCAGATGTCAGAGCCACAGACAGTTGCG 436

```

```

114 egluglnhrleuufyrphelysaspglyPro...LeuasmnleleuA 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 ATGATGCGGTGATGCTACACATCCATCCACAGGCTTGAATTCGAA 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 SPASPLGLYSLASPLEUHRASPLEUHRASPLEUHRASPLEUHRASPLE 139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
487 ACAGCGGCTGTCTCCAGAGCCACAGAGAGAGCTGACAACTCGAATACG 536
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 ..... 139
537 CTGCGTTCTTGTACACTCAAGCAGGCTTTCACCAAGACAGAACACT 586
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 ..... 139
587 GGCACACA.....GTGCTGCGGACATGAACGGTGTTCGAA 624
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
140 ...HISTRLYSTYRPROGLNLEUENUPROGLYLEAGGLYIESERLU 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
625 GAGACAAACAAGTGTCCACCGCTTACAGCTCGAAGAGAGGCA 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 GLUHRTHRLHGLYALHISASPLEUHRYSMETETALASNGLYIL 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
675 ACTCTCTTCCCGCATCAACGTCACAGAGCGCTTACAAAGTCCAGT 724
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 ELEULYSVALPROALALIASNVALASNASPSERVALTHRYLSERLYSP 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
725 TCGATACATCTACGCGTCCGCGCATCTTATCGATGATCAACGT 774
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 HNASPSNLEUHRYSGLYASARGLUSERLEULIASPGLYILEYARG 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
775 GCTTCGATGTCATGATCGCGGCGCAAGACACTCTGTCATGCTTACG 824
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 ALATHRASPYALMETILEALGLYLSVALALVALALAGLYTLEI 222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
825 CGATGTCGCAAGGCGTCCGCTCAATCCCTCGTGGCAAGCGCTCGG 874
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 YASPYALGLYLSGLYSSALAGLALALALALALALALALALALALAL 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
875 TTATCATACAGAACTGACCAATCTGCGCTCTCCAGCTTCGATGAA 924
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 ALLELIERHGLULIASPROIENALALALALALALALALALALALAL 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
925 GCGTACAGTCCCGCATGAGAACTGCGAAGTGTGATATCTT 974
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 GLYTRGLUVALTHRLHMETASPLUALACYSGLNGLUGLYASNIHEH 272
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
975 CGTACATGCGACAGAACTCGATATCATCTCTGTGACATGATGCC 1024
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 EVALTHRLTHRLGLYSILIASPILIELEUGLYARGLHSPHEGLUG 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 AGATGAAGATTAAGCTATTGCTGAACATGCGCACTTCGATACGAA 1074
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 LMETLYASPSAPALALALVALCYASNIIEGLYHISHPASPYALGLU 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1075 ATTGATACAGATGGCGCTCATGAAATACCGCATCAACAGATCCCAT 1124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 ILESPVALLYSTRPLEUASNGLU...ASNALVALGLULYSVALASNI 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1125 CAAGCCAGAAATACAGATGTGGAAATCCACAGATGCCACGATCTCC 1174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321 ELYSPROGLNVALASPARGYRARGLEULYSASNGLYARGLILIELE 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1175 TTCTTGTGAGAGCGCGCTTCTTAACCTTGGCGCTACAGTCAACCA 1224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
338 EULEUHLAIGLUGLYARGLEUVALASNLEUGLYCYSLALMETGLYHIS 354
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1225 TCTTTCGTTATGTCATGCTTCAACAAACAGACACTGCTGAGCTCGA 1274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 SERPHEVALMETSERASNSERHERHTRASNGLIVALMETALAGLNI 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1275 CCTTACGAAAAAGAGAAATCTCGAAGAAAGATTTACACACTTCGGA 1324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Align seg 1/1 to: US-08-896-005-3 from: 1 to: 504

34 GAGTACAGAAATTCCGGCAGCAATCCCAATCGTATCTCGCCGCTAGGAAGT 83
 67 AsphnCysValLysSerLleSerLysSerAlaPheLysAlaGlyValI 83
 84 TACCTTTCCTGAGAGAAATGGCAGGCTTATATGGTCTTCGTAGAGGTT 133
 83 eglIleAlaGlnSerGluMetProGlyIleMetThrLeuArgLysAla 100
 134 ATTCGGCTTCATAGCCATTGAAAGGTTCCAGAAATCTGTGGTCCCTCCAC 18
 100 IaLysAspGluLysProLeuLysGlyAlaLysAlaIleValGlyCysThrHis 110
 184 ATGCAGCTCCAGACAGCCGCTCCATCCGAGAAACATCCAGATCTGGTGGTC 23
 117 ValAsnAlaGlnSerAlaValLeuIleGluThrLeuValGlnLeuGlyAla 133
 234 TCAATGTCAGATGGGCTTCTGCAACATCTTCTATACCAAGATACAGCCG 28
 133 adhThrValArgThrAlaAlaCysAsnIleTyrSerThrGlnAsnAlaVala 15
 284 CTGGTCTATTCGTTTCGTGGCCCAACAGGCACACACAGAGAACCCAGCGGT 33
 150 IaAlaAlaLeuAla.....GluAlaGly 157
 334 ATCCAGTCTTGGCTGGAGAGGCGAAGAACATCCAGAAATATGCGAGAA 38
 156 IleProIlePheMetAlaTyrArgLysIuThrGluGluIuThrPtryc 174
 384 CACATACGCGCGCTCCACATGGCCCAATGGTCAAGGCCCAACAGAGTTG 43
 174 sLeuAspArgAlaIle...TyrSerAspGlyTyrGlnProAsnLeuIle 190
 434 TCGATGATGGTGGTGAATGCTATACACTCTCATCTCCAGAGGCTTCGAATTC 48
 190 euAspAspArgLysAlaThrHisLeuMetLeuLysTyr..... 204
 484 GAATAAGCGCGGTGTCTTCCAAAGCCAAAGAAAGCTGACAACTCGAATA 53
 205ProAspTyrPheLysAla..... 210
 534 CCGGCTGCTTTCGTAACATCAGACAGGCTTTCACCAAGACAGAAC 58
 210 210
 584 ACTGCCACACACTTGTGCGCGCATGCAACGCTGTTCCAGAGACAGACA 63
 211IleArgGlyIleValGlnGlnSerVal 219
 634 ACAGGTGTCACACGCGCTTACACAGGCTCGAGAGAGGGAAGCTCTCTT 66
 220 ThrGlyValHisArgLeuIuMetLeuSerLysGlyLysLeuThrVala 236
 684 CCCAGCCATCAACGTCACAGACGCTTTCACAAAGTCCCAAGTTCGATACACA 73
 236 lProAlaIleAsnValaLysAspSerValThrLysAsnLysPheAspThrP 253
 734 TCTTACGGCTGGCGGCACACCCCTTATCGATGAGGTATCAACCGCTTCCGAT 78
 253 heTyrThrCysTyrAspSerIleLeuAspSerLeuLysArgThrThrAsp 269
 784 GTCATGATCGCGCGCAACACAGCTCTGTCATAGGGTTACAGCGATTCGG 83
 270 IleMetPheGlyGlyLysGlnValaValIleCysGlyTyrGlyAspAlaI 286
 834 CAAGGCGTGGGCTGAATCCCTCCGTCGGCCCAAGGCGCTCGCTTATCATCA 88
 286 YLysGlyCysLysAlaGlnSerLeuLysGlyGlnIleCysLysLeuValTyrValT 303
 884 CAGAACTCGACCAATCTGCGCTCTCCAGGCGCATGGAAGCTACACG 93
 303 hrgIuValAspProIleCysAlaLeuGlnAlaIleMetAspLysPheArg 319
 334 GTCCGCGCATCGAGAAAGTGTCAAGAGATGCGATATCTTCGTATCATG 983

```

|||||.....|||||.....|||||.....|||||.....
320 ValValaIgrdeAsnGluValIleArgThrValAspValValValThrAl 336
984 CACAGAAAGTGGCATATCATCTGTGTGACATGATGGCCAGATGAAGG 1033
336 aThcGlyAsnGlyAsnValIleThrArgAspHisMetAsnArgMetGlySA 353
1034 ATAAAGCATATGTGGTAAACATCGGCACCTGATGATACGAAATGTGACA 1083
353 sngGlyCysIleLeuGlyCysAsnMetGlyHisSerGlyIleAspVal 369
1084 GATGGCCCTCATGAAATATCCAGCATCAAGACATCCCA..... 1122
370 AsnGlyLeu.....HisThrProGluLeuThr 379
1123 .....ATCAAGCCAGAAATAGACATGTGGGAATCCAGATGGCCACG 1165
379 pGluArgValArgSerGlnValAspHisIleArgTyrProAspGlyArgm 396
1166 CTATGCTGCTTGTGTGAGGGCCGCTTCTTAACCTTGCGTGGCTACA 1215
396 eTlleIleLeuValaGluGlyArgLeuValaHisLeuSerGlySerThr 412
1216 GGTCAACCATCTTGTGTATGTCAATGTCAATTCACAAACAGACACTGCG 1265
413 ...IleSerSerPheValValSerValAlaSerSerThrGlnAlaVal 428
1266 TCAGCTCGACCTCTACGAAAGAGAGAAATCTCGAAGAAAGGTTTACA 1315
428 AleuIleGluLeuPheSerAlaProGlyArgTyrTyrSerAspValTyrL 445
1316 CACTTCGCAAGCATCTCGATGAAGAAGTCGCTGCCCTCCACCTCGCATCT 1365
445 euleuProGlySylsMetAspGluTyrValAlaSerLeuHisLeuAlaThr 461
1366 CTCGATGTCACCTTACAAAGCTTACACAGAGAGGCTGACTACATCAA 1415
462 PheAspAlaHisLeuThrGluLeuThrAspGluGlnSerTyrPheMetG 478
1416 GCTTCAGATTGAGGGGCTTACAAAGTCTGATGCTTACCGT 1455
478 yLeuAsnGlySylsGluProPheGlySylsAlaAsnTyrTyrArg 491
seq_name: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:us-08-896-005-1
seq_documentation_block:
: Sequence 1, Application us/08896005
: Patent No. 5854023
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corley, Neil C.
: APPLICANT: Lal, Preeti
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ FOR Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/896,005
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:

```

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0337 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 500 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BLADYUT04
: CLONE: 1519044
: US-08-896-005-1

alignment_scores:
Quality: 1033.00 Length: 473
Ratio: 3.038 Gaps: 7
Percent Similarity: 71.882 Percent Identity: 44.397

alignment_block:
US-09-759-990-1 x US-08-896-005-1 ..

Align seg 1/1 to: US-08-896-005-1 from: 1 to: 500

43 ATTGCCGACATCAACCTCCATGTTCTGGCCGTAAAGAACTTACCTGTC 92
...:.....:.....:.....:.....:.....:.....:
77 ValLysAsnIleGlySylsGlnAlaGluPheGlyArgGluIleGlnIleAl 93
93 TGAGAAAGAAATGCCAGGTCTTATGTTCTTCTGAGCGTTATTCGGCTT 142
93 aGluGlnAspMetSerAlaLeuIleSerLeuArgSylsArgAlaGlnGly 110
143 CTAAGCATTTGAAGGTTGTCAATCTGTTCCCTCCACATGACAGTC 192
110 IuysProlLeuAlaGlyAlaLysIleValGlyCysThrHisIleThrAla 126
193 CAGACAGCCGTCCTCATCGAGACACTCACAGCTCTGTTGCTGATGTCAG 242
127 GlnThrAlaValIleIleGluThrLeuGlySylsAlaLeuGlnCysArg 143
243 ATGGGCTTCTGCAACATCTTCTTACACAGATACAGCCGCTGCTGTA 292
143 gTyrSerAlaCysAsnIleTyrSerThrGlnAsnGluValAlaAlaAla 160
293 TCGTTGTGCGCCCAACAGGCACACAGAGAAAGCCAGCGGTATGCCAGTC 342
160 eula.....GluAlaGlyValAlaVal 167
343 TTCGCTGGAAGGCGGAACACATCCAGAAATACCTGGGAGAAACATACCG 392
168 PheAlaTyrPylsGlySerGlnAspAspThrPylsThrPylsAspArg 184
393 CGCTCTCATATGGCCAGATGTCGAAGCCACAGAGGCTTGTGATGATG 442
184 gCysValAsnMet...AspGlyTyrGlnAlaAsnMetIleLeuAspAsp 200
443 GTGGTATGTCATCACTCTCCATCTCCAGAGGCTTGAATTCGAACAGCC 492
200 IyGlyAspLeuThrHisTyrPylsCysSylsTyr..... 211
493 GGTGCTGTCCAGAGCCCAAGAAAGCTGCAACCTCGAATACCGTGGCT 542
211 ..... 211
543 TCTGTCTACACTCAAGCAGGTTTCAACCAAGAAAGAAACCACTGGCACA 592
212 .....ProAsnValIleLys..... 217

```

```

593 CAGTTCGTGGCGGATGATGACGCTGTTCGAGAGACACACACAGCTGTC 642
218 .....IIeArgglyIleValIeGluIleuSerValThrGlyVal 229
643 CACCGCTCTACAGCTCGAAGAGAGGCAACTCTCTCCAGCAT 692
230 HIsArgIleuTyrGlnIleuSerIysAlaGlyIleuGlyValProAlaMe 246
693 CACGCTCAACGAGCTGTTCACAAAGTCCAGTTCAGTACATCTACGGCT 742
246 IasnValIasnAspSerValThrIysGlnIysPheAsnIleuTyrCysC 263
743 GCGCGCACCTCCCTTATCGATGATGATGATGATGATGATGATGATGATG 792
263 ysArgGlnSerIleIleuAspGlyIleuIysArgThrThrAspValMetPhe 279
793 GCGCGCAAGACAGCTCTGTCATGGGTACGGCATGTCGAGTCCGAGGGCTG 842
280 GlyGlyIysGlnValValIValCysGlyTyrGlyGlyValIleGlyGlyC 296
843 CGCTCAATCCCTCGGTGGCGGAGGCGCTGCGCTTATCATACAGAACTCG 892
296 scyAlaIalaIalaIeulysAlaIeulGlyAlaIeValTyrIleThrGlnIleA 313
893 ACCCAATCTGCGCTCTCCAGGCTGCCATGGAAGGCTACAGAGTCCGCGCG 942
313 sprProIleCysAlaIeulGlnIalacysMetIaspGlyPheArgValIalIys 329
943 ATGAGGAGAGCTGTCACAGAGTTCATATCTTCGTTACATGACAGAGAA 992
330 IeulAsnIeulValIleArgGlnValAspValValIleThrCysThrGlyAs 346
993 CTGGGATATCATCTCTGTGACATGATGCGCCAGATGAAGATGAAGGCTA 1042
346 IlysAsnValIalThrArgGlnIleuIleuAspArgMetIysAsnSerCysI 363
1043 TTGTCGCTAACATGCGCCACTTCGATTAAGAAATGATACAGATGCGCTC 1092
363 IeValIysAsnMetGlyHisSerAsnThrGlnIleAspValThrSerIeu 379
1093 ATGAAATACCCAGCATCAGACATCCCAATCAAGCCAGATACGACAT 1142
380 ...ArgThrProGluIleuThrTTPGluArgValArgSerGlnValAspH 395
1143 GTGGGAATCCAGATGCGCAGCTATCCCTCTTCTGAGGGCGCGC 1192
395 sValIleIleTTPProAspGlyIysArgValIalIeulIeulAlaGlyI 412
1193 TTCTTAACTTGGCTGGCTACAGAGTCAACCATCTTTCGTTATGTCAAATG 1242
412 eulIeulAsnIeulSerCysSerThr...ValProThrPheValIeulSerIle 427
1243 TCATTTCACAAACGACACTGCTGACGCTGACCTTACAGAA...AAGAG 1289
428 ThrIalThrThrGlnAlaIeulAlaIeulIleGluIeulTyrAsnAlaIProG 444
1290 AGGAATCTCGAGAAAGAGTTTACACACTTCGAGACATCTGATGAAG 1339
444 uGlyArgTyrIysGlnAspValTyrIleuIleuProIysMetIaspGlyut 461
1340 AAGTCGCTCGCTCCAGCTCGATCTCTGATGCGCAGCTTACAAAGCTT 1389
461 yValAlaIleSerIeulIeulIeulProSerPheAspAlaIleIleThrGluIeu 477
1390 ACACAGAAAGAGCTGACTACATCAAGTTCAGTGAAGGCTCTTACAA 1439
478 ThrAspAspGlnAlaIleIysTyrIleuGlyIleuAsnIysAsnGlyProPheI 494
1440 GTCTGATGCTTACCGTTAT 1458
494 sProAsnTyrTyrArgTyr 500

```

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-930-894-4

```

seq_documentation_block:
; Sequence 4, Application US/08930894
; Patent No. 6037524
; GENERAL INFORMATION:
; APPLICANT: GREENLAND, Andrew James
; APPLICANT: DRAPER, John
; APPLICANT: SKIPSEY, Marc
; APPLICANT: WARNER, Simon
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,894
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00882
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9507381.3
; FILING DATE: 10-APR-1995
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: ARABIDOPSIS SHH PROTEIN
; US-08-930-894-4

alignment_scores:
Quality: 351.00 Length: 148
Ratio: 3.375 Gaps: 3
Percent Similarity: 70.270 Percent Identity: 51.351

alignment_block:
US-09-759-990-1 x US-08-930-894-4 ..
Align seg 1/1 to: US-08-930-894-4 from: 1 to: 140

280 GCGCGCTGCTGATGCTGTGCGCCCAACAGGACACACAGAGCCAGC 329
|||||.....:
2 AlalaIalaIalaIleAlaIeAlaAspSerAlaIaIa.....12
330 CGGTATCCAGTCTTGGCTGGAAGGGGAAACACTCCCAAGATACTGG 379
|||||.....:
13 .....ValPheAlaTTPGlyGlyIleuThrIleuGlnIleuTTP 26
380 AGAACACATACCGCGCTCTACATGCGCAGATGTCAGAGCCACAGAG 429
|||.....:
26 TPCysThrGluArgAlaIeulAspTTPGlyGlyIleuTTPProAsp 42
430 GTTGTGATGATGCTGATGCTACACTCTCT...ATCTCCAAAGGCTT 476
|||||.....:
43 IleValAspAspGlyGlyAspAlaIleuThrIleuThrArgIleuGlnIleuVal 59
477 CGAA.....TTGCAAAACAGCGGCTGCTGCTTCCAGAGCCAAAG 514
|||.....:

```


MEDIUM TYPE: Floppy disk

[illegible]

seq_documentation_block:

```

; Sequence 7, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Wilcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-204-740-7

```

alignment_scores:

```

Quality: 287.00 Length: 76
Ratio: 4.100 Gaps: 0
Percent Similarity: 92.105 Percent Identity: 72.368

```

alignment_block:

```
US-09-759-990-1 x US-08-204-740-7 ..
```

```
Align seg 1/1 to: US-08-204-740-7 from: 1 to: 97
```

```

37 TACAGAAATGCGACATCAACCTCCATGTTCTGCGCGCTAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
22 TyrLysValAlaAspIleGlyLeuAlaIatpGlyArgGlyAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
87 CCTTGCTGAGAGGAATGCGAGCTTATGTTCTTCTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
38 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 CCGCTTCTAAGCAATGGAAGGCTGTCAGAAATCTCTGTTCCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 erAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
187 ACAGTCCAGACAGCGCTCCTCATGAGACACTACAGCTCTTGCTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72 ThrValGluThrValAlaValIleGluThrLeuValThrLeuGlyAlaGlu 88
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 TGTGAGATGGGCTTCTGCAACATCTTC 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

88 uValGlnTrpSerSerCysAsnIlePhe 97

```
seq_name: /cgn2_6/prodata/2/iaa/6A_COMB.pep:US-09-081-167A-7
```

seq_documentation_block:

```

; Sequence 7, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083745nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-167A-7

```

alignment_scores:

```

Quality: 287.00 Length: 76
Ratio: 4.100 Gaps: 0
Percent Similarity: 92.105 Percent Identity: 72.368

```

alignment_block:

```
US-09-759-990-1 x US-09-081-167A-7 ..
```

```
Align seg 1/1 to: US-09-081-167A-7 from: 1 to: 97
```

```

37 TACAGAAATGCGACATCAACCTCCATGTTCTGCGCGCTAAGAACTTAC 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
22 TyrLysValAlaAspIleGlyLeuAlaIatpGlyArgGlyAlaLeuAs 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
87 CCTTGCTGAGAGGAATGCGAGCTTATGTTCTTCTGAGCGTTATT 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
38 pIleAlaGluAsnGluMetProGlyLeuMetArgMetArgGluArgTyrS 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 CCGCTTCTAAGCAATGGAAGGCTGTCAGAAATCTCTGTTCCCTCCACATG 186
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55 erAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
187 ACAGTCCAGACAGCGCTCCTCATGAGACACTACAGCTCTTGCTGCTGA 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```

87 CCTTGCTGAGAGAAATGCCAGGTCTTATGTTCTTGAGGCTTAT 136
   ::::::::::::::::::::::::::::::::::::::::::::::
38 pIleAlaGluAsnGluMetProGlyLeuMetArgGluArgGlyTyrS 55
   ::::::::::::::::::::::::::::::::::::::::::::::
137 CCGCTTCAAGCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCATG 186
   ::::::::::::::::::::::::::::::::::::::::::::::
55 eRAlaSerLysProLeuLysGlyAlaArgGlyLeaGlyCysLeuHisMet 71
   ::::::::::::::::::::::::::::::::::::::::::::::
187 ACAGTCCAGACGCGTCTCATCGAGACATCAGCTCTTGTCGCTGA 236
   ::::::::::::::::::::::::::::::::::::::::::::::
72 ThrValGluThrAlaValLeuIleGluThrLeuValThrLeuGlyAlaG 88
   ::::::::::::::::::::::::::::::::::::::::::::::
237 TGTCAGATGGGCTTCTGCAACATCTTC 264
   ::::::::::::::::::::::::::::::::::::::::::::::
88 uValGluTrpSerSerCysAsnIlePhe 97

```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep: PCT-US95-02521-7

```

seq_documentation_block:
; Sequence 7, Application PC/TUS9502521
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-02521-7

```

alignment_scores:

Quality:	287.00	Length:	76
Ratio:	4.100	Gaps:	0
Percent Similarity:	92.105	Percent Identity:	72.368

alignment_block:

US-09-759-990-1 x PCT-US95-02521-7 ..

Align seg 1/1 to: PCT-US95-02521-7 from: 1 to: 97

```

37 TACAGAAATGCCAGATCAACCTCATGTTCTGCGCGTAAGGAATTAC 86
   ::::::::::::::::::::::::::::::::::::::::::::::
22 TyrLysValAlaAspIleGlyLeuAlaAlaTrpGlyArgLysAlaLeuAs 38
   ::::::::::::::::::::::::::::::::::::::::::::::
87 CCGTTCGAGAGAAATGCCAGGTCTTATGTTCTTCTGAGCGTTAT 136
   ::::::::::::::::::::::::::::::::::::::::::::::
38 pIleAlaGluAsnGluMetProGlyLeuMetArgGluArgGlyTyrS 55
   ::::::::::::::::::::::::::::::::::::::::::::::
137 CCGCTTCAAGCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCATG 186
   ::::::::::::::::::::::::::::::::::::::::::::::
55 eRAlaSerLysProLeuLysGlyAlaArgGlyLeaGlyCysLeuHisMet 71
   ::::::::::::::::::::::::::::::::::::::::::::::
187 ACAGTCCAGACGCGTCTCATCGAGACATCAGCTCTTGTCGCTGA 236
   ::::::::::::::::::::::::::::::::::::::::::::::
72 ThrValGluThrAlaValLeuIleGluThrLeuValThrLeuGlyAlaG 88
   ::::::::::::::::::::::::::::::::::::::::::::::
237 TGTCAGATGGGCTTCTGCAACATCTTC 264
   ::::::::::::::::::::::::::::::::::::::::::::::
88 uValGluTrpSerSerCysAsnIlePhe 97

```

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: US-08-204-740-9

```

seq_documentation_block:
; Sequence 9, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allgretti & Wilcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 312-715-1000
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-204-740-9

```

alignment_scores:

Quality:	271.00	Length:	76
Ratio: <th>3.985</th> <td>Gaps: <td>0</td> </td>	3.985	Gaps: <td>0</td>	0
Percent Similarity: <td>89.474</td> <td>Percent Identity: <td>69.737</td> </td>	89.474	Percent Identity: <td>69.737</td>	69.737

alignment_block:

US-09-759-990-1 x US-08-204-740-9 ..

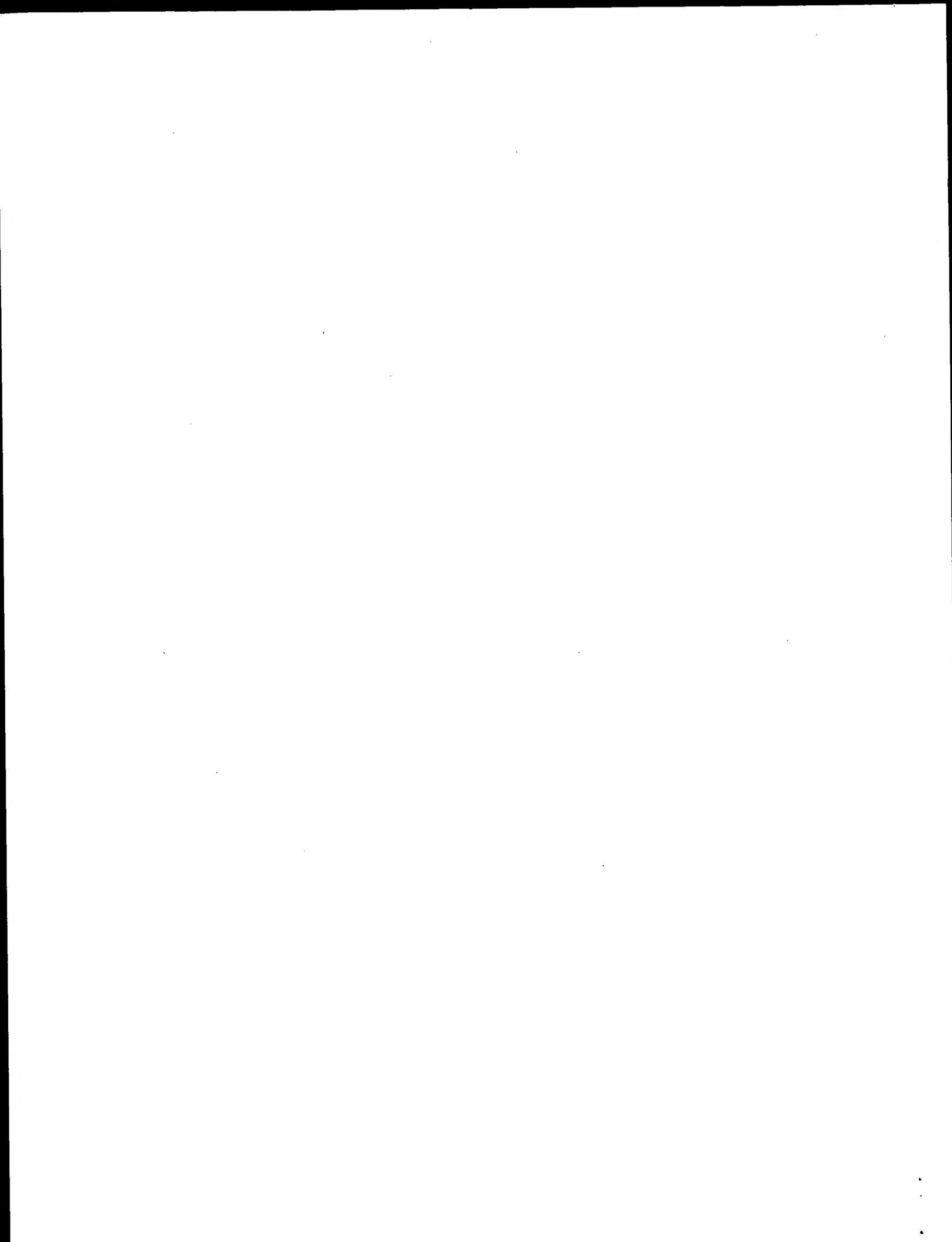
Align seg 1/1 to: US-08-204-740-9 from: 1 to: 95

```

37 TACAGAAATGCCAGATCAACCTCATGTTCTGCGCGTAAGGAATTAC 86
   ::::::::::::::::::::::::::::::::::::::::::::::
20 TyrLysValAlaAspIleGlyLeuAlaAlaTrpGlyArgLysAlaLeuAs 36
   ::::::::::::::::::::::::::::::::::::::::::::::
87 CCGTTCGAGAGAAATGCCAGGTCTTATGTTCTTCTGAGCGTTAT 136
   ::::::::::::::::::::::::::::::::::::::::::::::
36 pIleAlaGluAsnGluMetProGlyLeuMetArgGluArgGlyTyrS 53
   ::::::::::::::::::::::::::::::::::::::::::::::
137 CCGCTTCAAGCATTGAGGGGTGTCAGAAATCTGTGTTCCCTCCATG 186
   ::::::::::::::::::::::::::::::::::::::::::::::
53 eRAlaSerLysProLeuLysGlyAlaArgGlyLeaGlyCysLeuArgMet 69
   ::::::::::::::::::::::::::::::::::::::::::::::
187 ACAGTCCAGACGCGTCTCATCGAGACATCAGCTCTTGTCGCTGA 236
   ::::::::::::::::::::::::::::::::::::::::::::::
70 ThrValGluThrAlaValLeuIleGluThrLysValAlaLeuGlyAlaG 86

```


237 TGTTCAGATGGGCTTCCTGCAACATCTTC 264
:::|||||:::|||||
86 uAlaArgTrpSerSerCysAsnIlePhe 95



OM of: US-09-759-990-1 to: SwissProt_40:* out_format : pfs

Date: Aug 17, 2002 1:35 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=framer_n2p.model -DEV=slp
-O/cgn2_1/USPTO_spool/US0975990/runat_13082002_153926_28128/app_query.fasta_1.1541
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=n2p.rsp -CAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCI=0.000 -LOPEXT=0.000
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -LOPEXT=0.500
-DELOP=6.000 -RGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELEX=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -HEAPSIZ=500 -MTLEN=0 -MAXLEN=200000000
-USER=US0975990 @CGL1_1_39 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-759-990-1
Query length: 1461
Database: SwissProt_40:*
Database sequences: 105224
Database length: 38719550
Search time (sec): 49.660000
```

score_list:

```
Sequence Strd Orig zScore EScore Len
SwissProt_40:SAHH_TRIVA + 2534.00 3020.64 3.1e-161
SwissProt_40:SAHH_MTCU + 1550.50 1845.36 8.7e-96
SwissProt_40:SAHH_PETC + 1544.00 1837.77 2.4e-95
SwissProt_40:SAHH_ARATH + 1537.00 1829.41 6.9e-95
SwissProt_40:SAHH_TOBAC + 1533.00 1824.63 1.3e-94
SwissProt_40:SAHH_WHEAT + 1524.00 1813.88 5.0e-94
SwissProt_40:SAHH_CATRO + 1520.00 1809.10 9.3e-94
SwissProt_40:SAHH_MESCR + 1518.00 1806.71 1.3e-93
SwissProt_40:SAHH_PHASS + 1509.00 1795.95 5.0e-93
SwissProt_40:SAHH_MEDSA + 1499.00 1784.00 2.3e-92
SwissProt_40:SAHH_RHOCA + 1369.00 1629.14 1.0e-83
SwissProt_40:SAHH_RHOSH + 1359.00 1617.19 4.8e-83
SwissProt_40:SAHH_PLAF7 + 1332.00 1584.59 3.0e-81
SwissProt_40:SAHH_RAT + 1286.00 1530.67 3.4e-78
SwissProt_40:SAHH_MOUSE + 1278.00 1521.11 1.2e-77
SwissProt_40:SAHH_SCHPO + 1274.00 1516.29 2.1e-77
SwissProt_40:SAHH_HUMAN + 1260.00 1499.58 1.8e-76
SwissProt_40:SAHH_ANOGA + 1252.00 1490.62 5.7e-76
SwissProt_40:SAHH_DROME + 1251.00 1488.85 7.2e-76
SwissProt_40:SAH1_XENLA + 1242.50 1478.65 2.7e-75
SwissProt_40:SAH2_XENLA + 1238.50 1473.87 4.9e-75
SwissProt_40:SAHH_LEIDO + 1236.00 1470.79 7.2e-75
SwissProt_40:SAHH_DICDI + 1221.50 1453.63 6.6e-74
SwissProt_40:SAHH_YEAST + 1217.50 1448.42 1.2e-73
SwissProt_40:SAHH_CAEEL + 1203.50 1431.96 1.0e-72
SwissProt_40:SAHH_PNECA + 1142.00 1358.41 1.3e-68
SwissProt_40:SAH2_DROME + 1081.50 1284.78 1.4e-64
SwissProt_40:SAHH_SULSO + 1032.00 1225.72 2.8e-61
SwissProt_40:SAHH_AQUAE + 906.50 1077.56 6.1e-53
SwissProt_40:SAHH_METJA + 797.00 946.70 1.2e-45
SwissProt_40:SAHH_SYNX3 + 794.00 943.19 1.9e-45
SwissProt_40:SAHH_ARCFU + 781.00 927.42 1.4e-44
SwissProt_40:SAHH_PYRPU + 761.00 904.00 2.9e-43
SwissProt_40:SAHH_PYRAB + 739.50 877.93 7.9e-42
SwissProt_40:SAHH_PYRAB + 736.50 874.34 1.3e-41
SwissProt_40:SAHH_METTH + 735.50 873.12 1.5e-41
SwissProt_40:SAHH_THEMA + 734.50 872.05 1.7e-41
SwissProt_40:SAHH_THEMA + 696.50 826.96 5.7e-39
SwissProt_40:SAHH_STRFR + 372.00 451.24 1.6e-17
SwissProt_40:NSB2_YEAST + 206.50 229.88 3.2e-06
```

```
SwissProt_40:AGAL_YEAST + 202.00 230.33 5.4e-06 725 ! P32323 saccharomyces cere
SwissProt_40:YM96_YEAST + 200.50 224.06 7.7e-06 1140 ! Q04893 saccharomyces cer
SwissProt_40:YGIIF_YEAST + 190.50 219.30 2.9e-05 551 ! P53214 saccharomyces cere
SwissProt_40:YGLX_HSVB + 178.50 201.31 0.0002 757 ! P28968 equine herpesvirus
SwissProt_40:YHC8_YEAST + 177.50 202.84 0.0002 605 ! P38739 saccharomyces cere
```

seq_name: SwissProt_40:SAHH_TRIVA

seq_documentation_block:

```
ID SAHH_TRIVA STANDARD; PRT; 486 AA.
AC P51540;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalidea; Trichomonadida; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WAA38;
RX MEDLINE=97047381; PubMed=88923301;
RA Bagnara A.S., Tucker V.E., Minotto L., Howes E.R., Ko G.A., /
RA Edwards M.R., Davies I.W.;
RT "Molecular characterisation of adenosylhomocysteinease from
RT Trichomonas vaginalis."
RL Mol. Biochem. Parasitol. 81:1-12(1996).
CC -!- FUNCTION: ADENOSYLMHOCYSTEINE IS A COMPETITIVE INHIBITOR OF
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLMHOCYSTEINE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLMHOCYSTEINE.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -!- COFACTOR: NAD (BY SIMILARITY).
CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
CC -!- SIMILARITY: BELONGS TO THE ADENOSYLMHOCYSTEINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U40872; RAC47319.1; --
CC HSSP; P10760; IB3R.
CC InterPro; IPR000043; Ado_hcyase.
CC Pfam; PF00670; AdoHcyase; 1.
CC PROSITE; PS00738; AdoHcyase_1; 1.
CC PROSITE; PS00739; AdoHcyase_2; 1.
CC Hydrolase; NAD; One-carbon metabolism.
CC NP_BIND; 267 298 NAD (POTENTIAL).
CC SEQUENCE 486 AA; 53439 MW; 1403G6750587427C CRC64;
```

alignment_scores:

```
Quality: 2534.00 Length: 486
Ratio: 5.225 Gaps: 0
Percent Similarity: 99.794 Percent Identity: 98.971
```

alignment_block:

US-09-759-990-1 x SAHH_TRIVA ..

Align seg 1/1 to: SAHH_TRIVA from: 1 to: 486

```
1 ATGGCTTGCAATACCTACTGGTCTCTCCATTCGAGTACAGAAATGCCGA 50
|||||
1 MetAlaCysLysSerProAlaGlyAlaProPheGluTyArgIleAlaAs 17
```

51 CATCAACCTCCATGTTCTCGCCCGCTAAGGAAGTACCCCTTGCTGAGAAG 100

|||||
17 pIleAsnLeuHisValLeuGlyArgLysGluLeuThrLeuAlaGluLysG 34
101 AAATGCCAGGTCTTATGGTCTTCTCGTGGCGTATTCGGCTTCTTAAGCCA 150
34 luMetProGlyLeuMetValLeuArgGluArgTyrSerAlaSerLysPro 50
151 TTGAAGGGTCTCAAGATCTCTGGTTCCTCCATCATGACAGTCCACAGC 200
51 LeuLysGlyValArgIleSerGlySerLeuHisMetThrValGlnThrAl 67
201 CGTCCCTCATGACACACTCACAGTCTTGGTGGCTGATGTCAGATGGGCTT 250
67 aValLeuIleGluThrLeuThrAlaLeuGlyAlaAspValArgTTPAlaAs 84
251 CCGCAACATCTTCTTACACAAGATACAGCGCTGCTGCTATCGTTGTC 300
84 erCysAsnIlePheSerThrGlnAspThrAlaAlaAlaIleValVal 100
301 GGCCCAACAGCCACACAGAGAACGCCCGGTATCCAGTCTTCGCCCTG 350
101 GlyProThrGlyThrProGluLysProAlaGlyIleProValPheAlaTr 117
351 GAAGGGGGAACACTCCAGAAATACCTGGGAGAACACATACCGCGCTCTCA 400
117 pLysGlyGluThrLeuProGluTyrTyrGluAsnThrTyrArgAlaLeuT 134
401 CATGCCAGATGGTCAAGGCCACACAGAGTGTGTCGATGATGGTGTAT 450
134 hrTyrProAspGlyGlnGlyProGlnGlnValValAspAspGlyGlyAsp 150
451 GCTACACTCCCTCATCTCAAGGGCTTCGAATTCGAAACAGCGGTGCTGT 500
151 AlaThrLeuLeuIleSerLysGlyPheGluPheGluThrAlaGlyAlaVal 167
501 TCCAGAGCAACAGAGTGTCAACCTCGAATACCGCTGCTGCTTCTGCTA 550
167 lProGluProThrGluAlaAspAsnLeuGluTyrArgCysValLeuAlaT 184
551 CACTCAAGCAGGTCTTCAACCAAGCAACCAACCTGGCACACAGTGTCT 600
184 hrLeuLysGlnValPheAsnGlnAspLysAsnHisTyrPheThrValAla 200
601 CGCGCATGACGGTCTTCCGAAGACACAACACAGGTGTCCACCGCCT 650
201 AlaGlyMetAsnGlyValSerGluGluThrThrThrGlyValHisArgLe 217
651 CTACAGCTCGAAGAGGGGCAACTCTCTTCCAGCCATCAACCTCA 700
217 tyrGlnLeuGluLysGluGlyLysLeuLeuPheProAlaIleAsnVala 234
701 AGCAGCTGTTTCAAGTCCAAAGTTCGATAACATCTACGGCTGCGGCCAC 750
234 snAspAlaValThrLysSerLysPheAspAsnIleTyrGlyCysArgHis 250
751 TCCCTTATCGATGGTATCAACCGTGTCTCGATGTCATGATCGCGCCAA 800
251 SerLeuIleAspGlyIleAsnArgAlaSerAspValMetIleGlyGly 267
801 GACAGCTCTCGTATCGGTTCACGGCATGTCGCAAGGCTGCGCTCAAT 850
267 sThrAlaLeuValMetGlyTyrGlyAspValGlyLysGlyCysAlaGlns 284
851 CCTCGGTGCCAAGCGCTCGGCTTATCATCATCAAGACATCGACCCCAATC 900
284 erLeuArgGlyGlnGlyAlaArgValIleIleThrGluValAspProIle 300
901 TCGGCTCTCCAGCTGCCATGGAAGCTACAGTCCGCGCCATCGAGGA 950
301 CysAlaLeuGlnAlaValMetGluGlyTyrGlnValArgIleGluG 317
951 AGTCGTCAAGGATGTCGATATCTTCGTTACATGACACAGGAACATCGGATA 1000
|||||

317 uValValLysAspValAspIlePheValThrCysThrGlyAsnCysAspI 334
1001 TCATCTCTGTGACATGATGCCCGACATGAAGATAAGGCTATTTGTCGGT 1050
334 leIleSerValAspMetMetAlaGlnMetLysAspLysAlaIleValGly 350
1051 AACATCGGCCACTTCGATPACGAAATTCATACAGATGGCCTCATGAATA 1100
351 AsnIleGlyHisPheAspAsnGluIleAspThrAspGlyLeuMetLys 367
1101 CCCAGGATCAAGCACATCCCAATCAAGCCAGACATACGATGTGGGAAT 1150
367 rProGlyIleLysHisIleProIleLysProGluTyrAspMetIlePglup 384
1151 TCCAGATGGCCACGCTATCTCTCTTCTGCTGAGGGCGGCTTCTTAAC 1200
384 heProAspGlyHisAlaIleLeuLeuAlaGluGlyArgLeuLeuAsn 400
1201 CTTGGCTCGCTACAGTCAACCCATCTTTCGTTATGTCATGTCATTCAC 1250
401 LeuGlyCysAlaThrGlyHisProSerPheValMetSerMetSerPheTh 417
1251 AAACAGACACTCGCTCGCTCGACCTCTACGAAAGAGAGAGAAATCTCG 1300
417 rAsnGlnThrLeuAlaGlnLeuAspLeuTyrGluLysArgGlyAsnLeuG 434
1301 AGAGAAGGTTTACACACTTCGGAAGCATCTCGATGAAGAGTGGCTCGC 1350
434 luMetLysValTyrThrLeuProLysHisLeuAspGluGluValValArg 450
1351 CTCACCTCGGATCTCTCGATGTCACCTTACAAAGCTTACACAGAAGCA 1400
451 LeuHisLeuGlySerLeuAspValHisLeuThrLysLeuThrGlnLysG 467
1401 GCGTCACTACATCAAGTTCACCTGAGGCTGCTTACAAAGTCTGATGCTT 1450
467 nAlaAspTyrIleAsnValProValGluGlyProTyrLysSerAspAlaT 484
1451 ACCGTTAT 1458
484 yrArgTyr 486
seq_name: SwissProt_40:SAHH_MYCTU
seq_documentation_block:
ID SAHH_MYCTU STANDARD; PRT; 494 AA.
AC O08364;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
hydrolase) (AdoHcyase).
GN AHCT OR SAHH OR RV3248C OR MT3346 OR MTCY20B11.23C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OC NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala R.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]

1228 TTCGTTATGTCATATGTCATTCACAAACCCAGACACTCGCTCAGCTCGAACC 1277
1229 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 PheValMetSerAsnSerPheAlaAsnGluThrIleAlaGlnIleGlu 434
1278 CTACGAAGAAGAGGAAATCTCGAGAGCAAGGTTTACACACTTCGGAAGC 1327
1279 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 utpThrLysAsnAspGluTyrAspAsnGluValTyrArgLeuProLysH 451
1328 ATCTCGATCAAGAAGTCGCCTCGGCTCCACCTCGCATCTCTCGATGTCAC 1377
1329 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 isLeuAspGluLysValAlaArgIleHisValGluAlaLeuGlyGlyHis 467
1378 CTTACAAGCTTACAGAGACAGCGGTGACTACATCAACGTTCCAGTTCA 1427
1379 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
468 LeuThrLysLeuThrLysGluGlnAlaGluTyrLeuGlyValAspValGI 484
1428 GGGTCCTTACAAGTCTGATGTTACCGGTTAT 1458
484 uGlyProTyrLysProAspHisTyrArgTyr 494

seq name: swissprot_40:SAHH_PETCR

seq documentation block:

```

ID SEQUENCE DOCUMENTATION_DOCUMENT
1 SAHH_PETCR STANDARD; PRT; 485 AA.
AC Q01781;
AC 01-APR-1993 (Rel. 25, Created)
AC 01-OCT-1993 (Rel. 27, Last sequence update)
AC 16-OCT-2001 (Rel. 40, Last annotation update)
DT DT
DT Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
DE SAHH OR SHH.
GN GN
GN Petroselinum crispum (Parsley) (Petroselinum hortense).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID:4043;
XX [1]
XX SEQUENCE FROM N.A.
XX TISSUE=Leaf;
XX MEDLINE=92262510; PubMed=1374911;
XX Kwallack P., Plesch G., Hahlbrock K., Somssich I.E.;
XX "Induction by fungal elicitor of S-adenosyl-L-methionine synthetase
XX and S-adenosyl-L-homocysteine hydrolase mRNAs in cultured cells and
XX leaves of Petroselinum crispum.;"
XX Proc. Natl. Acad. Sci. U.S.A. 89:4713-4717(1992).
XX [2]
XX SEQUENCE OF 259-485 FROM N.A.
XX Somssich I.E., Bollmann J., Hahlbrock K., Kombrink E., Schulz W.;
XX "Differential early activation of defense-related genes in
XX elicitor-treated parsley cells.;"
XX Plant Mol. Biol. 12:227-234(1989).
XX [-] FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
XX S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
XX THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
XX CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
XX CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
XX [-] CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
XX + L-homocysteine.
XX [-] COFACTOR: NAD.
XX [-] PATHWAY: ACTIVATED METHYL CYCLE.
XX [-] SUBUNIT: HOMOTETRAMER.
XX [-] TISSUE SPECIFICITY: MAINLY IN FLORAL BUDS AND STEMS.
XX [-] SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
XX -----
XX This SWISS-PROT entry is copyright. It is produced through a collabora-
XX tion between the Swiss Institute of Bioinformatics and the EMBL outstat-
XX ion at the European Bioinformatics Institute. There are no restrictions on
XX use by non-profit institutions as long as its content is in no way
XX modified and this statement is not removed. Usage by and for commer-
XX cial entities requires a license agreement (See http://www.isb-sib.ch/annou-
XX or send an email to license@isb-sib.ch).
XX -----
XX EMBL: M81885; AAA33856.1; -.
XX DR

```

```

198 ArgLeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTy 214
654 CCAGCTCGAAGAGGAGGCAAACTCCTCTCCAGGCATCAACGTCACAG 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
214 rGlnMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnA 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 ACGCTGTACAAAGTCCAAAGTTCGATACATCATACGGTCCGCCACTCC 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
231 spSerValThrLysSerLysPheAspAsnLeuTyrgLysArgHisSer 247
754 CTTATCATCGTATCAACCGCTGCTCCGATGTCATGATCGGGCGCAAGAC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
248 LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVa 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
804 AGCTCTCGTCATGGGTTACGGCGATGTCGGCAAGGCTCGCTCAATCC 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
264 lAlaLeuIleAlaGlyTyrgLysAspValGlyLysGlyCysAlaAlaAla 281
854 TCGGTGCCAAGCGCTCGCTTATCATCATACAGAACTCGACCAATCTGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
281 etLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 GCTCTCCAGGTCGCATGGAAGGCTACAGGTCCTCGCGCATCGAGGAAGT 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
298 AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuProLeuGluAspVa 314
954 CPTCAAGATGTCATATCTTCGTTACATGCACAGCAAACTCGCATATCA 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
314 lValSerGluValAspIlePheValThrThrThrGlyAsnLysAspIleI 331
1004 TCTCTGTGTGATGATGCCAGATGAAGGTAAGGCTATTTCGGTATAC 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
331 leMetValSerAspMetArgLysMetLysAsnAlaIleValCysAsn 347
1054 ATGGCCACTTCGATACGAATATATACAGATGGCTCATCAAAATACCC 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
348 lIleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyPr 364
1104 AGGCATCAACACATCCCATCAAGCCAGATACAGATGCGGAATGCC 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
364 oGlyValLysArgIleThrIleLysProGlnThrAspArgTrpValPheP 381
1154 CAGAT...GCCACGCTATCCTCTCTTGTGTGAGGCGCGCTCTTAAAC 1200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
381 roAspThrGlyArgGlyIleIleIleLeuAlaGluGlyArgLeuMetAsn 397
1201 CTGTGCTGCTACAGGTACCCATCCTTTCGTTATGTCAATGTCAATCAC 1250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
398 LeuGlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheTh 414
1251 AAACACGACACTCGCTCAGTCGACCTCTAGGAAAGAGA.....GGAA 1294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
414 rAsnGlnValIleAlaGlnLeuGluLeuIleuTrpAsnGluLysSerGlyL 431
1295 ATCTCGAAGAGAGGTTTACACACTTCGAGCATCTCGATGAAGAGTC 1344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
431 ystyrGluLysLysValTyValLeuProLysHisLeuAspGluLysVal 447
1345 GCTCGCTCCACTCGATCTCGATCTCGATCTCCACTTACAAAGCTTACAC 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
448 AlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuSerLy 464
1395 GAAGCAGGCTGACTACATCAACGCTTCAGTTGAGGGTCCCTTACAGTCTG 1444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
464 sAspGlnAlaAspTyIleSerValProValGluGlyProTyryLysProA 481
1445 ATGCTTACGCTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
481 lAhIstyrArgTy 485

```

seq_name: SwissProt_40:SAHH_ARATH

seq_documentation_block:

```

ID AC 023255; 081847; STANDARD; PRT; 485 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHCyase).
GN SAHH OR AT4G13940 OR DL3010W.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERRECTA;
RA Belbahri L., Elleuch H., Villarreal R., Inze D., Thomas D.,
RA Thomasset B.;
RT "The isolation of an Arabidopsis thaliana cDNA clone encoding S-
RT adenosyl-L-homocysteine hydrolase";
RL (In) Plant Gene Register PGR99-139.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
RA Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T.M.,
RA Terry N., Giesen J., Villarreal R., de Clerck R., van Montagu M.,
RA Lecharny A., Auborg S., Gy I., Kreis M., Lao N., Kavanagh T.,
RA Hempel S., Kottler P., Entian K.-D., Rieger M., Schaeffer M., Funk B.,
RA Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,
RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatopoulou P.,
RA Piravandi E., Obermaier B., Hilbert H., Dueterhoeft A., Moores T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechman S., Ansoorge W.,
RA Cooke R., Berger C., Delsen M., Voet M., Volckaert G., Mewes H.-W.,
RA Klosterman S., Schueller C., Chaiwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana";
RL Nature 391:485-488(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Dueterhoeft A., Stiekema W., Entian K.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holzer B., Brandt A., Peters S., van Staveren M., Dircke W.,
RA Mooljman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernheiser S., Hempel S., Feldpausch M., Lamberth S., van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarreal R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloecker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedfor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,

```

us-09-759-990-1.n2p.rsp

[illegible]

[illegible]

106ValPheLatrPlpLysGlyGluThrLeuGlnGluTyrTirPcp 120
384 CACATACCGCGCTCTACATGCCAGATGGTCAAGGCCACACAGAGGTG 433
120 sThrGluArgAlaLeuAspTrpGlyProGlyGlyProAspLeuIleV 137
434 TCATGATGGTGGTATGCTACACTCTCTCATCTCCAAAGGGTTC..... 477
137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153
478GAATTCGAAACAGCGCGTCTCTCCAGAGCCCAACAGAGCTGA 521
154 GluGluGluPheAlaLysAsnGlyThrIleProAspProAsnSerThrAs 170
522 CAACCTCGAATACCGCTGCTGTCTGTCTACACTCAAGAGGTCTTCAACC 571
170 pAsnAlaGluPheGlnLeuValLeuThrIleIleLysGluSerLeuLysT 187
572 AAGACAAGAACCTGGCCACACAGTGTCTGCCGGCATGAACGCTGTTCC 621
187 hrAspProLeuLysTyrThrLysMetLysGluArgLeuValGlyValSer 203
622 GAACAGACAACAACAGGTGTCCACCGCTCTTACCAGCTCGACAAGAGGG 671
204 GluGluThrThrThrGlyValLysArgLeuTyrGlnMetGlnAlaAsnG 220
672 CAACACTCTCTCCAGGCATCAACGCTCAACAGCAGCTGTACAAATCCA 721
220 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 237
722 AGTTCGATAACACTACCGCTGCCGCGCACTCTTATCATGATGTATCAAC 771
237 ysPheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGCTCCGATGTATGATGTCGGCGCAAGACAGACTCTCTGTCATGGTTA 821
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaLeuValAlaGlyT 270
822 CGGCATCTCGCAAGGCTGGCTCAATCCCTCGCTGGCGCAAGGCGCTC 871
270 rGlyAspValGlyLysGlyCysAlaAlaLeuLysGlnAlaGlyAlaA 287
872 GCGTATCATCAGAACATCGACCCCAATCTGCGCTCTCCAGGTCGCCATG 921
287 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaThrMet 303
922 GAAGCTACCAAGTCCGCGCATCGAGGAGTCTCAAGGATGTCGATAT 971
304 GluGlyLeuGlnValLeuThrLeuGluAspValValSerAspValAspI 320
972 CTTCGTTATCATGCACAGAACTCGCATATCATCTCTGTTGACATGATGG 1021
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValAspHisMet 337
1022 CCAGATGAAGATAAGGCTATTGTCGGTAACTCGCCACTCGCCACTTCGATAAC 1071
337 rGlyAspMetLysAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
1072 GAAATGATACAGATCGCTCATGAAATACCCAGGCATCAAGCATCATCC 1121
354 GluIleAspMetLeuGluGluThrTyrProGlyValLysArgIleTh 370
1122 AATCAAGCCAGATACAGATCGGGAAATTCACAGATGGCCACGCT...A 1168
370 rIleLysProGlnThrAspArgTrpValPheProAspThrAsnSerGlyI 387
1169 TCCTCTCTTCTGCTAGGCGCGCTTCTTAACTTGGCTGCTGCTACAGGT 1218
387 leIleValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCCATCTTCTGTTATGCTCAATGCTATTACAAACACAGACACTCGCTCA 1268

seq_name: SwissProt_40:SAHH_WHEAT

seq_documentation_block:
ID SAHH_WHEAT STANDARD: PRT: 485 AA.

AC P32112:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase).
DE SAHH OR SHH OR SH6.2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Richards K.D., Gardner R.C.;
RL Submitted (XX-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS; THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine + L-homocysteine.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: ACTIVATED METHYL CYCLE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; L11872; AAA34303.1; --
DR HSPB; P10760; IB3R.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; ADHCHYASE_1; 1.
DR PROSITE; PS00739; ADHCHYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
NP_BIND 263 294 NAD (POTENTIAL).
FT SEQUENCE 485 AA; 53436 MW; 97FE208E3A6874ED CRC64;

alignment_scores:
Quality: 1524.00 Length: 487
Ratio: 3.868 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 61.807

alignment_block:

US-09-759-990-1 x SAHH_WHEAT ..

Align seg 1/1 to: SAHH_WHEAT from: 1 to: 485

```

19 ACTGGTCTCCATTCGAGTACAGAAATTCGGACATCAACCTCCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTyrLysValLysAspLeuPheGlnAlaAspPh 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCCGTAAAGAACTACCTTCGCTGAGAGAAATGCCAGGCTTATGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMet 41
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCTTCGTGACGGTATTCCGGCTTCTAAGCCATTGAAGGTGTGCAATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGGTTCCCTCCACATCACAGTCCAGACAGCGCTCTCATCGAGACACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGGTGTGATGTCAGATGGCTTCCTGCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerS 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAAGATACAGCGCTGCTGCTATCTGTTGTCGGCCCAACAGGCACACA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 erGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAGCAGCGGTATCCCACTCTTCCTCGGAGGCGGAACACTCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIlePheGlyGluThrLeuGlu 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATACCTGGAGACACATACCCTGCTCCATGTCGACAGTGTCTCAAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 uGluTyrTrpCysThrGluArgCysLeuAspTrpGlyValGlyGly 132
419 GCCCAGCAGCAGGTGTGATGATGGTGTGTGTGTGTGTGTGTGTGTGTGT 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 LysProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 148
469 AAGGGGTTC.....GAATTCGAACACGCGGTGTGTGTGTGTGTGTGT 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149 GluGlyValLysAlaGluGluPheGluLysSerGlyLysValProAs 165
507 GCCAAGCAGAGCTGACAACTCGAATACCGCTGCTGTCTTCTGTACATCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 ProGluSerThrAspAsnProGluPheLysIleValLeuThrIleLeA 182
557 AGCAGGTCTTCAACCAAGACAAGAACCACTGGCACACAGATTGCTGCCGGC 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 rgAspGlyLeuLysThrAspAlaSerLysTyrArgLysMetLysGluArg 198
607 ATGACCGGTGTTCCGAGAGACAACAACAGGTGTCCACCGCTCTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199 LeuValGlyValSerGluThrThrThrGlyValLysArgLeuTyrGlu 215
657 GCTCAGAGAGGAGGCAAACTCTCTTCCAGCCATCAAGTCAACAGCAG 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 nMetGlnGluSerGlyThrLeuLeuPheProAlaIleAsnValAsnAsp 232
707 CTGTTACAAAGTCCAAAGTTCGATATCTACCGCTGCCCGCTCCCTT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 exValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 248
757 ATCGATGTGTATCAACCGCTTCCGATGTCATGATCGCGGCGGCAAGCAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
807 TCTCGTATGGTTCAGCGGATGTGCGGCAAGGCTCGCTCAATCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 aValValCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaLeuLeu 282

```

```

857 GTGCCCAAGCGCTCGCTGTTATCATCATCAGAACTCGACCCCAATCTCGCT 906
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 ysglnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
907 CTCAGGCTGCCATGGAAGCTACAGGTCGCGCGCATCGAGGAAGTCTGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLeuGluAspValVa 315
957 CAAGGATGTCGATATCTCTGTTATCATCATCAGAGAACTCGCATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315 lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleLeu 332
1007 CNGTTCACATGATGGCCCATGAGGATAGGCTATTGTTCGGTAACATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 348
1057 GCCCACTTCGATAACGAAATGTATACAGATGGCTCATGAAATACCCAGG 1106
349 GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluThrTyrProG 365
1107 CATCAGCACATCCCAATCAAGCCAGAAATACGACATGTGGAAATCCCG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrValPheProG 382
1157 ATGGCCAC...GCTATCTCTCTTCTGCTGAGGCGCGCTTCTTAACTT 1203
382 luthrLysThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
1204 GGCTCGCTACAGGTACACCATCTTTCGTTATGTCATATGTCATTACAAA 1253
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
1254 CCAGACACTCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAAATC 1297
415 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysAlaSerGlyLys 432
1298 TCGAGAAGAAGGTTTACACACTTCCGAAGCATCTCGATGAAGAAGTCT 1347
432 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
1348 CGCTCCACCTCGGATCTCTCGATGTCACCTTACAAAGCTTACACAGAA 1397
449 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysSe 465
1398 GCAGGCTGACTACATCAAGTTCAGGTCAGGTCCTTACAGTCTGATG 1447
465 rGlnSerAspTyrIleSerIleProIleGluGlyProTyrLysLeuArgL 482
1448 CTWACCGTTAT 1458
   |||:|||||:
482 eutyArgTyr 485

```

seq_name: SwissProt_40:SAHH_CATRO

seq_documentation_block:

```

ID SAHH_CATRO STANDARD; PRT; 485 AA.
AC P35007;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-OCT-1994 (Rel. 30, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN SAHH.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vincet; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94218405; PubMed=8165255;

```

Schroeder G., Waitz A., Hotze M., Schroeder J.;
 *cDNA for S-adenosyl-L-homocysteine hydrolase from Catharanthus
 roseus*;
 Plant Physiol. 104:1099-1100(1994).
 -1- FUNCTION: ADENOSYLMHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
 S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 THEREFORE ADENOSYLMHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
 CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CONCENTRATION OF ADENOSYLMHOMOCYSTEINE.
 -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
 + L-homocysteine.
 -1- COFACTOR: NAD.
 -1- PATHWAY: ACTIVATED METHYL CYCLE.
 -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 -1- INDUCTION: BY STRESS.
 -1- SIMILARITY: BELONGS TO THE ADENOSYLMHOMOCYSTEINASE FAMILY.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)
 or send an email to license@isb-sib.ch.

 EMBL; 226981; CAA81527.1; -
 PIR; S38379; S38379.
 HSSP; P10760; 1B3R.
 InterPro; IPR000043; Ado_hcyase.
 InterPro; IPR000205; NAD_binding.
 Pfam; PF00670; Adohcyase; 1.
 PROSITE; PS00738; ADOHCYASE_1; 1.
 PROSITE; PS00739; ADOHCYASE_2; 1.
 Hydrolase; NAD: One-carbon metabolism.
 NAD (POTENTIAL).
 SEQ_ID 263 294
 SEQUENCE 485 AA; 50CC0E99A9F66C51 CRC64;
 SO

```

alignment_scores:      Length: 488
      Quality: 1520.00      Gaps: 6
      Ratio: 3.848
Percent Similarity: 80.943      Percent Identity: 62.295

alignment_block:
US-09-759-990-1 x SAHL_CATRO      ..

Align seg 1/1 to: SAHL_CATRO      from: 1 to: 485

19  ACTGGTGTCCATTCCGAGTACAGAAATTCCGACATCAACCTCCCATGTTCT 68
    |||:||||:|||||:||||:||||:||||:||||:||||:||||:||||:
8  ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24

69  CGGCCGTGAAGAACTTACCCTTGCTGTGAGAGGAATGCCAGGTCTTATGG 118
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
24  eGlyArgLeuGluIleGluLeuAlaGluValIgluMetProGlyLeuMetS 41

119 TTCTTCGTGAGCGTATTCCGCTTCTTAAGCCATTGAAGGGTGTCAGATC 168
    |||  :|||:||||:||||:||||:||||:||||:||||:||||:||||
41  erCysArgAlaGluPheGlyProSerGlnProPheLysGlyAlaLysIle 57

169 TCTGGTTCCTCCACATGCACAGTCCAGACAGCCGCTCCTCATCGACACACT 218
    :||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58  ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74

219 CACAGCTCTTGGTGCTGATGTGCAGATGGGCTTCCTGCACACATCTTCTCTA 268
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
74  uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerT 91

269 CACAAGATACAGCCGCTGCTGCTATCGTGTGTCGGCCCAACAGGCACACCA 318
    |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
91  hrGlnGluHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105

319 GAGAAGCCAGCGGTATCCCACTTTCGCTGTAAGGGCGGAACACACTCCC 368

```



```

807 TCTGTCATGGTGTACGGGATGTCGGCAAGGCGTGGCTCAATCCCTCC 856
      :::::|||||
265 yvalValCysGlyTyrGlyAspValGlyCysAlaLeuAlaLeuL 282
      :::::|||||
857 GTGGCAAGGCGCTCGGTTATCATCAGAACTCGACCACTCTCGCT 906
      :::::|||||
282 ysAlaAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
      :::::|||||
907 CTCAGGCTGCCATGGAAGCTACCAAGTACCGCGCATCGAGAACTCGT 956
      :::::|||||
299 LeuGlnAlaLeuMetGluGlyPheGlnIleLeuThrLeuGluAspValva 315
      :::::|||||
957 CAAGATGTCGATATCTGTTACATGCACAGAAACTGGATATCATCT 1006
      :::::|||||
315 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleIle 332
      :::::|||||
1007 CTGTTGACATGATGGCCAGATGAAGTATAGCTATTGCGGTAACTC 1056
      :::::|||||
332 etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
      :::::|||||
1057 GGCCACTTCGATAACCAAAATGATACAGATGGCTCATGAAATACCCAGG 1106
      :::::|||||
349 GlyHisPheAsnGlnIleAspMetLeuGlyLeuGluAsnTyrProG 365
      :::::|||||
1107 CATCAGCAGATCCCAATCAAGCCAGATACGATCTGGGAATTCACG 1156
      :::::|||||
365 yvalLysArgIleThrIleLysProGlnThrAspArgPheValPheProG 382
      :::::|||||
1157 ATGCCAC...GCTATCTCTCTTCTGCTGAGGCGCGCTTCTTAACCTT 1203
      :::::|||||
382 luthrAsnThrGlyIleIleValLeuAlaGluGlyArgLeuMetLysLeu 398
      :::::|||||
1204 GGCTGCGCTACAGTACCATCTTTCGTTATGTCATGATCAATGATTCACAAA 1253
      :::::|||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
      :::::|||||
1254 CCAGACATCGCTCAGCTCGACCTCTACGAAAGAGA...GGAAATC 1297
      :::::|||||
415 nglValIleAlaGlnLeuGluLeuTrpAsnGluArgAlaSerGlyLys 432
      :::::|||||
1298 TCGAGAAGAGTTTACACATCTCGAAGATCTCGATGAGAAGTCTGCT 1347
      :::::|||||
432 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 448
      :::::|||||
1348 CGCTCCACCTCGATCTCGATCTGATCTACACCTTACAAAGCTTACACAGAA 1397
      :::::|||||
449 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuSerLysAs 465
      :::::|||||
1398 GCAGGCTGACTACATCAAGCTTCCAGTTGAGGCTTACAGTCTGATG 1447
      :::::|||||
465 pGlnAlaAspTyrIleSerValProValGluGlyProTyrLysProAlaH 482
      :::::|||||
1448 CTTACCGTTAT 1458
      :::::|||||
482 lstyArgTyr 485
      :::::|||||

seq_name: SwissProt_40:SAHH_PHASS
seq_documentation_block:
ID: SAHH_PHASS STANDARD; PRT: 485 AA.
AC P50249;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN SAHH.
OS Phalaenopsis sp.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Epidendroideae; higher Epidendroideae; Vandaeae; Aeridinae;
OC Phalaenopsids.
```


334 ATCCAGTCTTCGCTTGGAGGCGGAAACACACTCCCAAGAACTACTGGGAGAA 383
 106ValPheAlaTrpLysGlyGluThrLeuGlnGluTrpTrpCy 120
 384 CACATACCGCGCTTCACATGCCAGATGTCAGGCGCCACAGCAGGTTG 433
 120 sThrGluArgCysLeuGluTrpGlyAlaGlyGlyGlyProAspLeuIleV 137
 434 TCGATGATGGTGGTGATGCTACACTCCTCATCTCCCAAGGCGCTTC..... 477
 137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153
 478GAATTCGAAACAGCGGTGCTGCTTCACACTCAAGCAGGCTTC 521
 154 GluGluGluTrpGlyLysAsnGlyLysIleProAspProAlaSerThrAs 170
 522 CAACCTCGAATACCGCTGCTTCTGCTACACTCAAGCAGGCTTC 571
 170 pAsnAlaGluPheGlnIleValLeuGlyLeuIleArgAspSerLeuSerV 187
 572 AAGACAAGAACCACTGGCACACAGTGTGCTGCGCGGATGAACGGTGTTC 621
 187 alAspProLysLysTyArgMetLysGluArgLeuValGlyValSer 203
 622 GAAGACAGACAACAGGTGTCCACCGCTCTACACAGCTCGAAGAGG 671
 204 GluGluThrThrThrGlyValLysArgLeuTrpGlnMetGlnTrpSerG 220
 672 CAACCTCCTTCCAGCCATCAACGTCAAGCAGCTGTTACAAAGTCCA 721
 220 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 237
 722 AGTTCGATACACTACGCTGCTGCGCGCACTCCCTTATCGATGATCAAC 771
 237 ySphAspAsnLeuTrpGlyCysArgHisSerLeuProAspGlyLeuMet 253
 772 CGTGTCTCCGATCATGATGCTGCGCGGCAAGACAGCTCTGCTATGGGTTA 821
 254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270
 822 CGGCGATGTCGCAAGGCTGCGCTCAATCCTCGTGGCGCAAGCGCTC 871
 270 rGlyAspValGlyLeuGlyCysAlaAlaLeuLysThrAlaGlyAlaA 287
 872 GGTATCATCATCAGAACTCGACCAATCTGCGCTCTCCAGCTCCCATG 921
 287 rGValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
 922 GAAGCTACAGTCCGCGCGCATCGAGGAAGTCTCAAGGATGCGATAT 971
 304 GluGlyLeuProValLeuArgLeuGluAspValValSerGluAlaAspI 320
 972 CTTCGTTACATGCAGAGAACTGCGATATCATCTCTGTTGACATGATGG 1021
 320 ePheValThrThrThrGlyAsnLysAspIleIleMetValAspHisMet 337
 1022 CCAGATGAAGTAAGGCTATGTCGGTAACATCGGCCACTTCGATAAC 1071
 337 rGlyMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
 1072 GAAATGTATACAGTGGCTCATGAATACCCAGGATCAAGCACATCCC 1121
 354 GluIleAspMetLeuGlyLeuGluSerPheProGlyValLysArgIleTh 370
 1122 AATCAAGCAGAAATACGACATGTGGGAATTCACAGATGGCCAGCT...A 1168
 370 rIleLysProGlnThrAspArgArgValPheProAspThrAsnSerGlyI 387
 1169 TCCTCTCTTGTGCTGAGGCGCCGCTTCTTAACCTTGGCTGGCTACAGGT 1218
 387 leLeuValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403

1219 CACCCATCTTTCGTTATGTCATCTCATTCAACAACACAGACACTCGCTCA 1268
 404 HisProSerPheValMetSerSerPheThrAsnGlnValIleAlaG 420
 1269 GCTCGACCTCTAGCAAAAGAGA.....GGAAATCTCCAGAGAAGGTTT 1312
 420 nLeuGluLeuTrpLysGluArgAlaSerGlyLysTrpGluLysValT 437
 1313 ACACATTCGGAAGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGA 1362
 437 yValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453
 1363 TCTCTCGATGTCACCTTACAAAGCTTACACAGAAGCAGGCTGACTACAT 1412
 454 LysLeuGlyAlaLysLeuThrLysLeuThrProSerGlnAlaAspTy 470
 1413 CAACGTTCCAGTTGAGGCTCTTCAAGTCTGATGCTTACCGTTAT 1458
 470 eSerValProValGluGlyProTyLysProAlaHisTyArgTy 485
 seq_name: SwissProt_40:SAHH_MEDSA
 seq_documentation_block:
 ID SAHH_MEDSA STANDARD; PRT; 485 AA.
 AC P30246;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Adenosylhomocysteinease (EC 3.3.1.1) (S-adenosyl-L-homocysteine
 DE hydrolase) (AdoHcyase).
 GN SAHH.
 OS Medicago sativa (Alfalfa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 OX NCBI_TaxID=3879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SIRIVER; TISSUE=Stem;
 RX MEDLINE=95201245; PubMed=7894016;
 RA Abrahams S., Hayes C.M., Watson J.M.;
 RT "Expression patterns of three genes in the stem of lucerne (Medicago
 sativa).";
 RL Plant Mol. Biol. 27:513-528(1995).
 CC -!- FUNCTION: ADENOSYLHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
 CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
 CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O -> adenosine
 CC + L-homocysteine.
 CC -!- COFACTOR: NAD (BY SIMILARITY).
 CC -!- PATHWAY: ACTIVATED METHYL CYCLE.
 CC -!- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L36119; AAB41814.1; -.
 DR HSSP; P10760; 1B3R.
 DR InterPro; IPR000043; Ado_hcyase.
 DR Pfam; PF00670; AdoHcyase_1.
 DR PROSITE; PS00738; ADHOCYASE_1; 1.
 DR PROSITE; PS00739; ADHOCYASE_2; 1.
 KW Hydrolase; NAD; One-carbon metabolism.
 FT NP_BIND 263 294 NAD (POTENTIAL).
 FT SEQUENCE 485 AA; 53151 MW; 13EAFDBD527DF59F CRC64;

alignment_scores:
 Quality: 1499.00 Length: 482
 Ratio: 3.853 Gaps: 4
 Percent Similarity: 80.705 Percent Identity: 61.411

alignment_block:
 US-09-759-990-1 x SAHH_MEDSA ..

Align seg 1/1 to: SAHH_MEDSA from: 1 to: 485

```

34 GAGTACAGAAATCCGACATCAACCTCCATGTTCTCGCGCGCTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GlutyllysvallylaspmetserGlnAlaasphegylArgLeuGluI 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84 TACCCTTGCTGAGAAGAAATGCCAGTCTTATGGTCTTCGTCGAGCGTT 133
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 eGluLeuAlaGluValGluMetProGlyLeuMetSerCysArgThrGluP 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 ATCCCGCTTCTAAGCATTTGAAGGTCTCAAGATCTCTGGTTCCTCCAC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
46 heglyProserGlnProPheylslyAlaArgilethrGlySerLeuHis 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 ATGACAGTCCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 TGATGTCAGATGGCTCTTCGCAACATCTTCTACACAAGATACAGCGG 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79 aGluValArgTrpCysSerCysAsnIlePheSerThrGlnaspHisAlaA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
284 CTGCTGCTATCTGTTGTCGCGCCACAGGACACACCAGAGAGCCAGCGGT 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
96 laAlaAlaIleAlaArgaspSerAlaAla..... 105
334 ATCCAGCTCTTCGCTGGAAGGGGAAACACTCCAGATATCTGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 .....ValPheAlaTrpLysGlyluThrLeuGlnGluTrpTrpCy 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
384 CACATACCGCTCTCACATGCCAGATGGTCAAGGCCACACAGAGGTG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
120 sserGluArgAlaLeuaspTrpGlyProGlyGlyProaspLeuIleV 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
434 TCGATGATGGTGGTGAATGCTACATCTTCATCTCAAGGGCTTCGAA... 480
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
137 alaspaspGlyGlyaspValThrLeuLeuIleHisGluGlyVallyAla 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
481 .....TTGAAACAGCGGTGCTGTTCAGAGCCCAACAGAGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 GluGluValPheGluLysThrGlyGlnLeuProaspProSerSerThrAs 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
522 CAACCTCGAATACCGCTCGTCTTGTGCTACACTCAAGCGGTCTCAACC 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
170 pAsnAlaGluMetGlnIleValLeuThrIleleArgaspGlyLeuLys 187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
572 AAGACAAGAACACTGGGACACAGTGTGTCGCGGCATGACGGTGTTC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
187 hrAspProlylserArgTrpGlnLysMetLysThrArgIleValGlyValser 203
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
622 GAAGACACAAACAGGTGTCACCGCTCTTACAGCTCGAGAGGAGGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
204 GluGluThrThrThrGlyVallylserArgLeuTyrglnMetglnAlaSerG 220
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
672 CAAACTCTCTTCCAGCGCATCAACGCTCAACGACGCTGTACAAAGTCCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
220 yThrLeuLeuPheProAlaIleAsnValAspSerValThrLysSerL 237
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
722 AGTTCGATACATCTACGCTCGCGCTCCCTCTATCGATGGTATCAAC 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
237 yspPheAspAsnLeuTyrglyCysArgHisSerLeuProaspGlyLeuMet 253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
772 CGTGTCTCGCATGTCATGTCGCGCGCAACAGACGCTCTCGTCATGGGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270
822 CGCGGATGTGCGCAAGGGCTCGCTCAATCCCTCGTGGCGCAAGCGCTC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
270 rGlyAspValGlyLysGlyCysAlaAlaAlaLeuLysGlnGlyAlaAla 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
872 GCCTTATCATCACAGAACTGACCCCAATCTCGCTCTCCAGGCTGCCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
287 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
922 GAAGGCTACAGGTCGCGCGCATCGAGGAAGTCGTCGAAGGATGCGATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 GluGlyLeuGlnValLeuThrIleGluAspValIleSerGluAlaAspI 320
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
972 CTTCGTTTACATGCACAGAAACTCGCATATCATCTCTGTGTGACATGATG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
320 ePheValThrThrThrGlyAsnLysAspIleIleMetValSerAspMetL 337
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1022 CCCAGATGAAGGATAAGCTATTGTCGGTAACTACATCGCCACTTCGATAAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
337 yLysMetLysAsnAlaIleValCysAsnIleGlyHisPheaspasn 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1072 GAAATTCATACAGATGGCTCATGAATACCCAGGCATCAAGCACATCCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 GluIleaspMetHisGlyLeuGluThrTrpProGlyValLysArgIleTh 370
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1122 AATCAAGCCAGAAATACGACATGTGGAAATCCAGATGCGCCAGT... 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
370 rIleLysProGlnThrAspArgTrpValPheProGluThrLysSerGlyI 387
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1169 TCCCTCTTCTGTCGAGGCGCCCTTCTTAACCTTGGCTGCGCTACAGGT 1218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 leIleValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1219 CACCATCTTCTGTTATGTCATATGTCATTCACAAACACAGACATCGCTCA 1268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaGl 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1269 GCTCGACCTCTACGAAAGAGA.....GGAATCTCGAGAAGAGGTTT 1312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
420 nIleGluLeuTrpLysGluLysThrSerGlyLysTyrglyLysValT 437
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1313 ACACACTTCCGAGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGA 1362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
437 yrValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1363 TCTCTCGATGTCACCTTACAAAGCTTACACAGAGCAGCGCTGACTACAT 1412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 GlnLeuGlyAlaLysLeuThrLysLeuSerLysAspGlnAlaAspTyrl 470
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1413 CAACGTTCCAGTTGAGGTCCTTACAAAGTCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
470 eSerValProValGluGlyProTyrlLysProAlaHisTyrglyTyr 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: SwissProt_40:SAHH_RHOCA

seq_documentation_block:
 ID SAHH_RHOCA STANDARD; PRT; 463 AA.
 AC P28183;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
 hydrolase) (AdoHcyase).
 GN AHCY.
 OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 OX NCBI_TaxID=1061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SB1003 / ST LOUIS;
 RX MEDLINE=92335291; PubMed=1631127;

```

alignment_scores:
  Quality: 1369.00
  Ratio: 3.670
  Percent Similarity: 77.066
  Length: 484
  Gaps: 8
  Percent Identity: 59.504

alignment_block:
  US-09-759-990-1 x SAHH_RHOCA
  ..
  Align seg 1/1 to: SAHH_RHOCA from: 1 to: 463

34 GAGTACAGAAATGCCGACATCAACCTCCATGTTCTCGCGCGTAGGAAC 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 AsptyrileValylasaspIleLysLeuAlaGluPheClyArgLysGluLe 19
84 TACCCTTGCTGAGAGGAATGCCAGGTCTTATGGTCTCTCGTACGGTT 133
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
19 uaspileAlaGluThrGluMetProGlyLeuMetAlaCysArgGluGluP 36
134 ATTCCGCTTCTAAGCATTGAGGGTGTGAGAATCTCTGGTTCCCTCCAC 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
36 heglyProserGlnProLeuLysGlyAlaArgIleAlaGlySerLeuHis 52
184 ATGACAGTCCACAGACGCCCTCCTCATCGACACACTCACAGCTCTTGGTGC 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 MetThrIleGluThrAlaValLeuIleGluThrLeuLysAlaLeuGlyAl 59
234 TGATGTCAGATGGGCTTCTCGACATCTTCTCTACACAAGATACAGCGG 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 aaspValargTrpAlaSerCysAsnIlePheSerThrGlnAspHisAlaA 86
284 CTGCTGCTATGTTTCGGCCCAACAGGCACACCACAGAGAAGCACGCCGT 333
   | : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
1207 TGGCTACAGGTCACCCATCTTTGTTATGTCATTCATTCACAAACCA 1256
|||||
380 AsnAlaThrGlyHisProSerPheValMetSerAlaSerPheThrAsnGI 396
|||||
1257 GACATCGCTCAGCTCGACCTCTACGAAAGAGAGAAATCTCGAGAAGA 1306
|||||
396 nValLeuAlaGlnIleGluLeuThrPheThrLysGlyAlaGluTyrGlnProg 413
|||||
1307 AGTTTACACATTCGAGACATCTCGATCAAGAGTTCGCTCGCTCCAC 1356
|||||
413 lyValTyrIleLeuProLysSerLeuAspGluLysValAlaArgLeuHis 429
|||||
1357 CTCGGATCTCTCGATCTCCACCTTACAAAGCTTACACAGAGAGGCTGA 1406
|||||
430 LeuLysIleGlyValLysLeuThrThrLeuArgProAspGlnAlaGI 446
|||||
1407 CTACATCAAGCTTCCAGTTCGAGGCTCTTACAGTCTGATGCTTACCGTT 1456
|||||
446 uTyrIleGlyValThrValGluGlyProPheLysSerAspHisTyrArgT 463
|||||
1457 AT 1458
||
463 yr 463
```

seq_name: SwissProt_40:SAHH_RHOSH

```
seq_documentation_block:
ID SAHH_RHOSH STANDARD; PRT; 463 AA.
AC 050562:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase).
GN AHCY.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=97354111; PubMed=9210332;
RA Mizoguchi H., Masuda T., Nishimura K., Shimada H., Ohta H., Shioi Y.,
RA Takamiya K.
RT "Nucleotide sequence and transcriptional analysis of the flanking
RT region of the gene (spb) for the trans-acting factor that controls
RT light-mediated expression of the put operon in Rhodobacter
RT sphaeroides."
RL Plant Cell Physiol. 38:558-567(1997).
CC -|- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -|- COPACTOR: NAD (BY SIMILARITY).
CC -|- PATHWAY: ACTIVATED METHYL CYCLE.
CC -|- SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U76671; AAB88245.1; -
DR HSP; P10760; IB3R.
DR InterPro; IPR000043; Ado_hcyase.
DR Pfam; PF00670; AdoHcyase; 1.
DR PROSITE; PS00738; ADHOCYASE_1; 1.
DR PROSITE; PS00739; ADHOCYASE_2; 1.
KW Hydrolase; NAD; One-carbon metabolism.
FT NP_BIND 247 278 NAD (POTENTIAL).
SEQUENCE 463 AA; 50617 MW; ACE8EB9D0D0CCBBA CRC64;

```
alignment_scores:
Quality: 1359.00 Length: 487
Ratio: 3.634 Gaps: 8
Percent Similarity: 76.797 Percent Identity: 58.932

alignment_block:
US-09-759-990-1 x SAHH_RHOSH ..
Align seg 1/1 to: SAHH_RHOSH from: 1 to: 463

34 GAGTACAGAAATGCGACATCAACCTCATGTTCTCGCGCGTAAGAACT 83
|||||
3 AspPheIleValLysAspLeuSerLeuAlaAspPheGlyArgLysGluLe 19
|||||
84 TACCTTCTGAGAGAGAAATGCCAGGTCTTATGGTTCTTCGTGACCGTT 133
|||||
19 uAspPheAlaGluThrGluMetProGlyLeuMetAlaLeuArgGluGluP 36
|||||
134 ATTCCGCTTCTAAGCCATTGAAGGGTGTGAGATCTCTGGTTCCTCCAC 183
|||||
36 heGlyAlaSerLysProLeuLysGlyAlaArgIleAlaGlySerLeuHis 52
|||||
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCACAGCTCTTGGTC 233
|||||
53 MetThrValGlnThrAlaValLeuIleGluThrLeuValAlaLeuGlyAl 69
|||||
234 TGATGTCAGATGGCTTCTCTGCAACATCTTCTACACAGATACACCGG 283
|||||
69 aspValArgTtpAlaSerCysAsnIlePheSerThrGlnAspHisAla 86
|||||
284 CTGCTCTATCTGTTGTCGCCCAACAGGCACACAGAGAAAGCCAGCCGGT 333
|||||
86 laAlaAlaIleAlaAla.....SerGly 93
|||||
334 ATCCCGATCTTCGCTGGAAGGCGAAACACTCCAGAAATACTCGGAGAA 383
|||||
94 ValProValPheAlaIleLysGlyGluThrLeuGluAspTyrTtpAlaTy 110
|||||
384 CACATACCGCGCTCTCATGCGCCAGATGGTCAAGGCCACAGAGGTTG 433
|||||
110 rThrAspLysIlePheGlnPheAlaAspGlyThr...CysAsnMetIleLe 126
|||||
434 TCAGATGATGGTGGTATGCTACACTCTCATCTCCAAAGGCTTCGAATTC 483
|||||
126 euAspAspGlyGlyAspAlaThrLeuTyrIleLeuLeuGlyAlaArgVal 142
|||||
484 GAACAGCCGGT.....GCTGTTCCAGAGCCCAACAGAGAAGC 518
|||||
143 Glu...AlaGlyGluThrAspLeuIleAlaValProGlnSerAspGluGI 158
|||||
519 TGCAACCTCGAATACCGCTCGGTCTTGTCTACACTCAAGCAGGCTTCA 568
|||||
158 uValCysLeu.....PheA 163
|||||
569 ACCAAGCAAGAACCCAC.....TGGCACACAGTTGCT 600
|||||
163 snGlnIleArgLysArgMetAlaGluThrProGlyTyrPheThrLysGln 179
|||||
601 GCC...GCATGAAGCGGTGTTCCGAAGAGACAAACAGGTGTCACCG 647
|||||
180 ArgAspAlaIleLysGlyValSerGluGluThrThrGlyValHisar 196
|||||
648 CCTCTACACAGTTCGAGAGAGGCGAAACTCTCTTCCAGCCGCTCAACG 697
|||||
196 gLeuTyrAspLeuHisLysLysGlyLeuLeuProPheProAlaIleAsnV 213
|||||
698 TCAACGACGCTGTTCACAAAGTCCAAAGTTCGATAACATCTACGCTGCCG 747
|||||
213 aLAsnAspSerValThrLysSerLysPheAspAsnLysTyrGlyCysLys 229
|||||
748 CACTCCCTTATCGATGATGATCAACCGTGTCTCCGATGTGATGATCGCGG 797
```


99 ValPheAlaTrpLysAsnGluThrLeuGluGluTyrTrpCysValG1 115
 390 CCGCGCTCCTACATGGCCAGATGTT CAAGGCCACACAGGAGTGG 433
 115 userAlaLeuThrTrpGlyAspGlyAspAsnGlyProAspMetIleVal 132
 434 TCGATGATGGTGGTGTACATCTCCTCATCTCCAAAGGGCTTCGAA... 480
 132 AlaAspGlyGlyAspAlaThrLeuLeuValHisGlyValGluTyr 148
 481TTCCGAACACCGCGTCTCTTCCAGAGCCACACAGAGCTGA 521
 149 GluLysLeuTyrGluGluLysAsnIleLeuProAspProGluLysAlaL 165
 522 CAACTCGAATACCGCTCGTCTTCTGTCTACACTCAAGCAGGCTTCAACC 571
 165 SasnGluGluGluArgCysPheLeuThrLeuLeuLysAsnSerIleLeuL 182
 572 AAGACAAGAACCACTGGCCACAGTGTGCTCGCGCATGACCGTGTTCCT 621
 182 YAsnProLysLysTrpThrAsnIleAlaLysLysIleIleGlyValSer 198
 622 GAAGAGACAACAGAGTGTCCACCGCTCTACACCTCGAGAGGAGG 671
 199 GluGluThrThrGlyValLeuArgLeuLysLysMetAspLysGlnAs 215
 672 CAAACTCTCTTCCAGCCATCAACGCTCAACGAGCTGTACAAAGTCCA 721
 215 nGluLeuLeuPheThrAlaIleAsnValAsnAspAlaValThrLysGlnL 232
 722 AGTTCGATACATCTACGCTGCCGCCACTCCCTTCATCGATGGTATCAAC 771
 232 YsyrAspAsnValTyrGlyCysArgHisSerTyrThrAspGlyLeuMet 248
 772 CGTCTCCGATGTCATGATCGCGCGCAGACAGCTCTCGTCATGGTTA 821
 249 ArgAlaThrAspPheLeuIleSerGlyLysIleValIleCysGlyTyr 265
 822 CGGCGATCGCGAAGGGTGGCTCAATCCCTCCGTCGCGCAAGGGGCTC 871
 265 rGlyAspValGlyLysGlyCysAlaSerSerMetLysGlyLeuGlyAlaA 282
 872 CGGTATCATCACAAGTTCGACCAATCTGGCTCTCAGGCTGCCATG 921
 282 rGValTyrIleThrGluIleAspProIleCysAlaIleGlnAlaValMet 298
 922 GAAGGCTACCGAGTCCGCGCATCGAGGAGTCTCAAGGATCTCGATAT 971
 299 GluGlyPheAsnValThrLeuAspGluIleValAspLysGlyAspPh 315
 972 CTTCGTTATCATCAGGAACTCGCATATCATCTCTGTGTGACATGATGG 1021
 315 ePheIleThrCysThrGlyAsnValAspValIleLysLeuGluHisLeuL 332
 1022 CCCAGATGAGGATAGGCTATGTGCGGTAACTCGGCCACTTCGATAAC 1071
 332 eULysMetLysAsnAlaValAlaValGlyAsnIleGlyHisPheAspAsp 348
 1072 GAATTTGATACAGTGGCTCTCATGAATACCCAGGCATCAAGCACATCCC 1121
 349 GluIleGlnValAsnGluLeuPheAsnTyrLysGlyIleHisIleGluAs 365
 1122 AATCAAGCCAGAAATCAGACATGTGGAAATCCAGATGGCCACGCTATCC 1171
 365 nValLysProGlnValAspArgIleThrLeuProAsnGlyAsnLysIleI 382
 1172 TCCTTCTTCTGAGGCGCCCTCTTAACTTCGCTGCGCTACAGGTCAC 1221
 382 IeValLeuAlaArgGlyArgLeuLeuAsnLeuGlyCysAlaThrGlyHis 398
 1222 CCATCTTCTGTTATGTCATTCATTCACAAACACACACTCGCTCAGCT 1271
 399 ProAlaPheValMetSerPheCysAsnGlnThrPheAlaGlnLe 415

1272 CGACTCTACGAAAGAGA.....GGAAATCTCGAGAAAGGTTTACA 1315
 415 uAspLeuTrpGlnAsnLysAspThrAsnLysTyrGluAsnLysValTyrL 432
 1316 CACTTCCGAAGCATCTCGATGAGAGAGTCTCGCTCCACCTCGGATCT 1365
 432 euLeuProLysHisLeuAspGluLysValAlaLeuTyrHisLeuLysLys 448
 1366 CTCGATGTCCACTTACAAAGCTTACACAGACAGGCTGACTACATCAA 1415
 449 LeuAsnAlaSerLeuThrGluLeuAspAsnGlnCysGlnPheLeuG1 465
 1416 GTTCCAGTTGAGGTCTTACAAAGTCTGATGCTTACCGTTAT 1458
 465 yValAsnLysSerGlyProPheLysSerAsnGluTyrArgTyr 479

seq_name: SwissProt_40:SAHH_RAT

seq_documentation_block:

ID SAHH_RAT STANDARD; PRT; 431 AA.
 AC P10760;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Adenosylhomocysteine (EC 3.3.1.1) (S-adenosyl-L-homocysteine
 hydrolase) (AdoHcyase).
 GN AHcy.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=87118240; PubMed=3027698;
 RA Ogawa H., Gomi T., Mueckler M.M., Fujioaka M., Backlund P.S. Jr.,
 RA Aksamit R.R., Unson C.G., Cantoni G.L.;
 RT "Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat
 liver as derived from the cDNA sequence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:719-723(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FTSCHER 344;
 RX MEDLINE=95262723; PubMed=7744082;
 RA Merta A., Aksamit R.N., Kasir J., Cantoni G.L.;
 RT "The gene and pseudogenes of rat S-adenosyl-L-homocysteine
 hydrolase.";
 RL Eur. J. Biochem. 229:575-582(1995).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC TISSUE=Liver;
 RX MEDLINE=99315209; PubMed=10387078;
 RA Hu Y., Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioaka M.,
 RA Takusagawa F.;
 RT "Crystal structure of S-adenosylhomocysteine hydrolase from rat
 liver.";
 RL Biochemistry 38:8323-8333(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY. (2.8 ANGSTROMS).
 RX MEDLINE=20493601; PubMed=10913437;
 RA Komoto J., Huang Y., Gomi T., Ogawa H., Takata Y., Fujioaka M.,
 RA Takusagawa F.;
 RT "Effects of site-directed mutagenesis on structure and function of
 recombinant rat liver S-adenosylhomocysteine hydrolase. Crystal
 structure of D244E mutant enzyme.";
 RL J. Biol. Chem. 275:32147-32156(2000).
 CC -1- FUNCTION: ADENOSYLOHOMOCYSTEINE IS A COMPETITIVE INHIBITOR OF
 S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
 THEREFORE ADENOSYLOHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
 CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
 CONCENTRATION OF ADENOSYLOHOMOCYSTEINE.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
 + L-homocysteine.

129 spAspGlyGlyAspLeuThrAsnLeulle. 138
487 ACAGCGGGTCTGTTCCAGAGCCCAACAGAAGCTGACAACTCGAATACGG 536
138 138
537 CTGCGTCTTGTCTACACTCAAGCAGGCTTCAACCAAGACAAGAACCACT 586
138 138
587 GCACACA GTTGCTCCGCGCATGAACGGTGTGTTCCGAA 624
139 ..HisThrLysHisProGlnLeuLeuSerGlyIleArgGlyIleSerGlu 154
625 GAGACAACAACAGGTCTCCACGCCCTTACCAGCTCGAGAGGAGGGCAA 674
155 GluThrThrGlyValHisAsnLeuTyrLysMetMetAlaAsnGlyIle 171
675 ACTCCTCTTCCAGCCATCAACGCTCAACGACGCTGTGTACAAAGTCCCAAGT 724
171 eLeuLysValProAlaIleAsnValAsnAspSerValThrLysSerLysP 188
725 TCGATAACATCTACGGCTCCGCACTCCCTTATCGATGATCAACCGT 774
188 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyIleLysArg 204
775 GTTCCGATGTCATCGCGCGCAAGACAGCTCTCGTCATGGTTCACGG 824
205 AlaThrAspValMetIleAlaGlyLysValAlaValAlaGlyTyrGI 221
825 CGATGTCGGCAAGGCTCGCTCAATTCCTCCGTGGCCAGGGCGCTCGCG 874
221 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArgV 238
875 TTATCATCACAGAATCTCACCAATCTGGCTCTCCAGCTGCCATGGAA 924
238 alIleIleThrGluLeaspProIleAsnAlaLeuGlnAlaAlaMetGlu 254
925 GGCTACAGGTCGCCGATCGAGGAAGTCGTCAAGGATGTCGATATCTT 974
255 GlyTyrGluValThrThrMetAspGluAlaCysLysGluGlyAsnIlePh 271
975 CGTTACATGCACAGAAACTCGGATATCATCTCTTGCATGATGGCC 1024
271 eValThrThrGlyCysValAspIleIleLeuGlyArgHisPheGluG 288
1025 AGATGAAGGATAAGCTATTGTCGGTAACATCGGCCACTTCGATAACGAA 1074
288 lnMetLysAspAspAlaIleValCysAsnIleGlyHisPheAspValGlu 304
1075 ATTGATACAGATGGCTCATGAAATACCCAGGCATCAAGCATCCCAAT 1124
305 IleAspValLysTrpLeuAsnGlu..AsnAlaValGluLysValAsnIle 320
1125 CAAGCCAGAATACAGATGTGGGAATTCACAGATGGCCACCTCTCTCC 1174
320 eLysProGlnValAspArgTyrLeuLeuLysAsnGlyHisArgIleIleL 337
1175 TTCCTGCTCAGGGCCGCTTCTTAACCTTGGCTCGCTACAGGTCACCCA 1224
337 euLeuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 353
1225 TCTTCTGTTATGTCATTCATTCACAAACCACACTCGCTCAGCTCGA 1274
354 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleGI 370
1275 CCTCTACGAAAAGAGGAAATCTCGAGAGAAGGTTTACACACTCTCCGA 1324
370 uLeuThrThrHisProAspLysTyrProValGlyValHisPheLeuProL 387
1325 AGCATCTCGATGAAGAATCGCTCGCTCGCTCACCTCGAATCTCTCATGTC 1374

205 AlaThrAspValMetIleAlaGlyLysValAlaValValAlaGlyTyrGI 221
825 CGATGTGGCAAGGCCTCGCTCAATCCCTCGTGCCCAAGGCGTCGCG 874
221 YASPVALGLYSLGYCYSLAALNALAEUARGGLYPHEGLYALAARGV 238
875 TTATCATCACAAGACTCGACCCAAATCTGGCTCTCCAGGCTGCCATGGAA 924
238 alIleIleThrGluLeuLeuAspProIleAsnAlaLeuGlnAlaAlaMetGlu 254
925 GGCTACCAGGTCCCGCGCATCGAGGAAGTCGCAAGGATGTCGATATCTT 974
255 GlyTyrGluValThrThrMetAspGluAlaCysLysGluGlyAsnIlePh 271
975 CGTTACATGCACAGGAAGACTCGGATATCATCTCTGTTGACATGATGGCCC 1024
271 eValThrThrGlyCysValAspIleIleLeuGlyArgHisPheGlu 288
1025 AGATGAAGGATAAGGCTATTTCGGTAACATCGGCACCTTCGATAACGAA 1074
288 InMetLysAspAspAlaIleValCysAsnIleGlyHisPheAspValGlu 304
1075 ATTGATACAGATGGCTCATGAAATATCCAGGATCAAGCACATCCCCAAT 1124
305 IleAspValLysTrpLeuAsnGlu...AsnAlaValGluLysValAsnII 320
1125 CAAGCCAGAATACGACATGTGGGAATTCAGAGTGGCCACGCTATCCCTCC 1174
320 eLysProGInValAspArgTyrLeuLeuLysAsnGlyHisArgIleIleL 337
1175 TPTCTGTCTCAGGGCCGCTTCTTAACTTCGGCTCGCTACAGGTACCCCA 1224
337 euLeuAlaGluGlyArgLeuValAsnLeuGlyCysAlaMetGlyHisPro 353
1225 TCCTTTCGTTATGTCATTCACAAAACGACACTCGCTCAGCTCGA 1274
354 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleGl 370
1275 CCTCTACGAAAAAGAGGAAATCTCGAAGAAGGTTTACACACTTCCGA 1324
370 uLeuTrpThrHisProAspLysTyrProValGlyValHisPheLeuProL 387
1325 AGCATCTCGATGAAGAAGTTCGGCTCGGCTCCACCTCGGATCTCTCGATGTC 1374

Mon Aug 19 15:16:21 2002

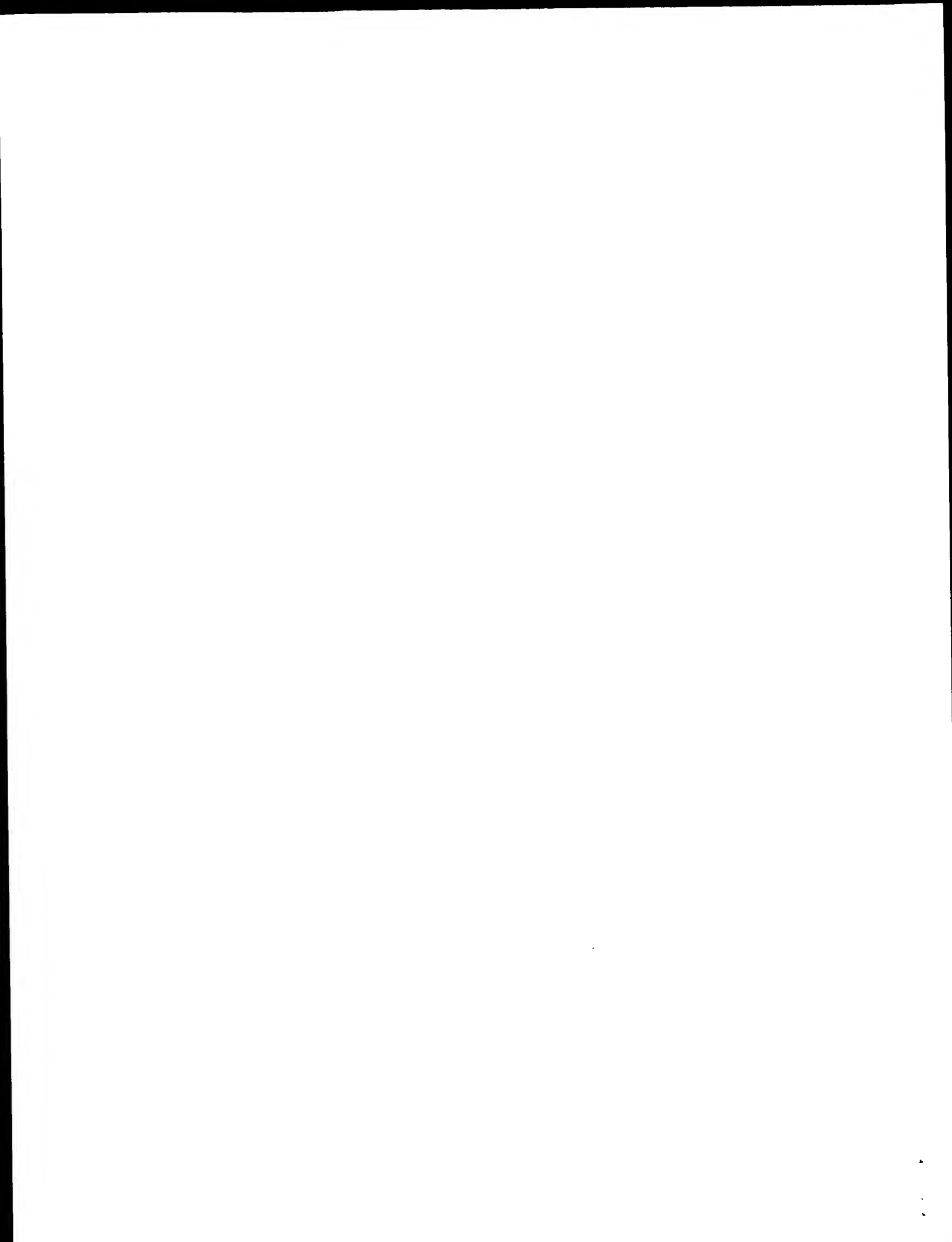
us-09-759-990-1.n2p.rsp

```

387 ysLysLeuAspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 403
1375 CACCTTCAAGGCTTACACAGAGGCTGACTACATCAACAGTTCCTCCAGT 1424
|||||
404 LysLeuThrLysLeuGluGlnAlaGlnTyrLeuGlyMetProI 420
1425 TCAGGGTCTTACAGTCTGATGCTTACCGTTAT 1458
420 eAsnGlyProPheLysProAspHisTyrArgTyr 431
seq_name: SwissProt_40:SAHH_MOUSE
seq_documentation_block:
ID SAHH_MOUSE STANDARD; PRT; 431 AA.
AC P30247; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine
DE hydrolase) (AdoHcyase) (Liver copper binding protein) (CUBP).
GN AHCY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Liver;
RX MEDLINE=95386522; PubMed=7657650;
RA Bethin K.E., Petrovic N., Rittinger M.J.;
RT Identification of a major hepatic copper binding protein as S-
RT adenosylhomocysteine hydrolase.
RL J. Biol. Chem. 270:20698-20702(1995).
CC S-ADENOSYL-L-METHIONINE-DEPENDENT METHYL TRANSFERASE REACTIONS;
CC THEREFORE ADENOSYLHOMOCYSTEINASE MAY PLAY A KEY ROLE IN THE
CC CONTROL OF METHYLATIONS VIA REGULATION OF THE INTRACELLULAR
CC CONCENTRATION OF ADENOSYLHOMOCYSTEINE.
CC -| CATALYTIC ACTIVITY: S-adenosyl-L-homocysteine + H(2)O = adenosine
CC + L-homocysteine.
CC -| COFACTOR: NAD.
CC -| PATHWAY: ACTIVATED METHYL CYCLE.
CC -| SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -| SUBCELLULAR LOCATION: Cytoplasmic.
CC -| SIMILARITY: BELONGS TO THE ADENOSYLHOMOCYSTEINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L32836; AAA70378.1; -
CC DR HSP; P10760; IB3R.
CC DR MGD; MGI:87968; Ahcy.
CC DR InterPro; IPR000043; Ado_hcyase.
CC DR Pfam; PF00670; AdoHcyase; 1.
CC DR PROSITE; PS00738; ADOHCYASE_1; 1.
CC DR PROSITE; PS00739; ADOHCYASE_2; 1.
CC KW Hydrolase; NAD; One-carbon metabolism.
CC FT INIT_MET 0
CC FT NP_BIND 213 244 NAD (POTENTIAL).
CC SEQUENCE 431 AA; 47543 MW; D57DCE9D98C9B5 CRC64;
alignment_scores:
Quality: 1278.00 Length: 478
Ratio: 3.662 Gaps: 5
Percent Similarity: 73.013 Percent Identity: 56.485
alignment_block:
US-09-759-990-1 x SAHH_MOUSE
Align seg 1/1 to: SAHH_MOUSE from: 1 to: 431
37 TACAGAAATCCCGACATCAACCTCCATGTTCTCGCGGTAGGAACCTAC 86
|||||
6 TyrLysValAlaAspIleGlyLeuAlaAlaTrpGlyArgLysAlaLeuAs 22
87 CTTCTGAGAGAAATGCCAGGTCTTATGTTCTTCTCGTGGCGTTATT 136
|||||
22 PileAlaGluAsnGluMetProGlyLeuMetArgMetArgGluMetTyrS 39
137 CGCTTCTAAGCCATTGAAGGTGTCAAGATCTCTGTTTCCCTCCACATG 186
|||||
39 eAlaSerLysProLeuLysGlyAlaArgIleAlaGlyCysLeuHisMet 55
187 ACAGTCAGACAGCGCTCTCATCGAGACACTCACAGCTCTTGGTGTGA 236
|||||
56 ThrValGluThrAlaValLeuIleGluThrLeuValAlaLeuGlyAlaGl 72
237 TGTCAAGTGGGCTCTCTGCAACATCTTCTTACACAAGATACAGCGCTG 286
|||||
72 uValArgTrpSerCysAsnIlePheSerThrGlnAspHisAlaAla 89
287 CTGCTATCGTTGTCGCCCAACAGCACACAGAGCCAGCCGCGTATC 336
|||||
89 laAlaIleAla.....LysAlaGlyIle 96
337 CCAGTCTTCGCTGGAAGGGCAACACTCCAGATATCTGGGAGAACAC 386
|||||
97 ProValPheAlaTrpLysGlyGluThrAspGluLysLeuTrpCysI 113
387 ATACCGGCTCTCATGATGCCAGATGTTCAAGCCCAACAGCAGCGTTGTCG 436
|||||
113 eGluGlnThrLeuHisPheLysAspGlyPro....LeuAsnMetIleLeu 129
437 ATGAGTGGTGGTGTGATGTCATCTCTTCCAAAGGCTTCGAATTCGAA 486
|||||
129 sPaspGlyLysAspLeuThrAsnLeuIle..... 138
487 ACAGCGGCTGCTTCCAGAGCCCAACAGAGCTGACACCTCGNATACCG 536
138 ..... 138
537 CTGCGTTCITGCTACACTCAAGCAGGTCTTCAACCAAGAACCAACT 586
138 ..... 138
587 GGCACACA.....GTTGCTCCCGCATGAACGGTGTTCGAA 624
|||||
139 ..HisThrLysTyrProGlnLeuSerGlyIleArgGlyIleSerGlu 154
625 GAGACAAACAGTGTCCACCGCTCTACAGCTCGAGAGAGGAGGCAA 674
|||||
155 GluThrThrThrGlyValHisAsnLeuTyrLysMetSerAsnGlyI 171
675 ACTCCTCTTCCCGCATCAACGTCAACGCTGTTCACAAAGTCCAAGT 724
|||||
171 eLeuAsnValProAlaIleAsnValAsnAspSerValThrLysSerLysP 188
725 TCGATACATCTACGGTGTCCCGCATCTCCCTTATCGATGGTATCAACGT 774
|||||
188 heAspAsnLeuTyrGlyCysArgGluSerLeuIleAspGlyIleLysArg 204
775 GCTTCCGATGTGATGATGGCGGCAAGACAGCTCTCGTATGGGTTACGG 824
|||||
205 AlaThrAspValMetIleAlaGlyLysValAlaValAlaGlyTyrGl 221
825 CGATGTCGGCAAGGCTCGCTCAATCTCCGTGGCCCAAGGCGCTCGCG 874
|||||
221 yAspValGlyLysGlyCysAlaGlnAlaLeuArgGlyPheGlyAlaArgV 238
875 TTATCATCACAGAACTCGACCAATCTGCGCTCTCCAGGCTGCCATGGAA 924

```


|||||
238 aillelethrGlulileasProIleasnalaleuGlnAlaIleMetGlu 254
925 GCCTACCGAGTCGCGCGCATCGAGGAGTCGTCAAGGATGTCGATATCTT 974
|||||
255 GlyTyrGluValThrThrMetaspGluAlaCysLysGluGlyAsnIlePh 271
975 CGTTACATGCACAGAACTGCCATATCATCTCTGTTGCACATGATGGCCC 1024
|||||
271 evalthrThrGlycysvalaspIleleLeuGlyArgHisPheGluG 288
1025 AGATGAAGGATGAAGCTATTGTGGTAACATCGGCCACTTCGATAACGAA 1074
|||||
288 InMetLysAspAlaIleValCysAsnIleGlyHisPheAspValGlu 304
1075 ATTGATACAGATGGCCTCATGAATACCGAGGCATCAAGCACATCCCAAT 1124
|||||
305 IleaspValLysTrpLeuAsnGlu...AsnAlaValGluLysValAsnI 320
1125 CAAGCCAGAATACGACATGTCGGAATTCACAGATGCCACGCTATCCTCC 1174
|||||
320 elysProGlnValaspArgTyrTrpLeuLysasnGlyArgIleIleL 337
1175 TTCTTGTGAGGCGCGCTTCTTAACCTTGGTCGCTACAGGTCACCCCA 1224
|||||
337 euLeuAlaGluGlyArgLeuValasnLeuGlyCysAlaMetGlyHisPro 353
1225 TCTTTCGTATGTCATGTCATTCACAAACGACACACTGCTCAGCTCGA 1274
|||||
354 SerPheValMetSerAsnSerPheThrAsnGlnValMetAlaGlnIleGl 370
1275 CCTCTACGAAAGAGAGGAAATCTCGAGAAGAAGTTTACACACTTCCGA 1324
|||||
370 uLeuTrpThrHisProAspLysTyrProValGlyValHisPheLeuProL 387
1325 AGCATCTCGATGAGAACTGCTCGCTCCGCTCCACCTCGGATCTCTCGATGC 1374
|| |||||
387 ysLysLeuaspGluAlaValAlaGluAlaHisLeuGlyLysLeuAsnVal 403
1375 CACCTTACAAAGCTTACACAGAGGAGGCTGACTACATCAACGTTCCAGT 1424
|||||
404 LysLeuThrLysLeuThrGluLysGlnAlaGlnTyrLeuGlyMetProIl 420
1425 TGAGGGTCCTTACAGTCTGATGCTTACCGTTAT 1458
:|||||
420 eAsnGlyProPheLysProAspHisTyrArgTyr 431



Sequence	Strd Orig	ZScore	EScore	Len	Documentation
ir2:B70593	+	1550.50	1828.11	495	adenosylhomocysteine
ir2:D87005	+	1537.50	1812.79	1,4e-94	adenosylhomocysteine
ir2:C71400	+	1537.00	1812.33	1,4e-93	probable S-adenosyl-L-homocysteine
ir2:T06764	+	1524.00	1796.94	1,5e-93	adenosylhomocysteine
ir2:S38379	+	1520.00	1792.21	1,1e-92	adenosylhomocysteine
ir2:S71621	+	1509.00	1779.19	2,0e-92	adenosylhomocysteine
ir2:D82730	+	1406.00	1658.08	1,1e-91	adenosylhomocysteine
ir2:D97362	+	1382.00	1629.25	6,5e-85	adenosylhomocysteine
ir2:F25F80	+	1382.00	1629.25	2,5e-83	adenosylhomocysteine
ir2:H87280	+	1374.00	1619.84	8,4e-83	adenosylhomocysteine
ir1:A46035	+	1352.50	1594.42	4,2e-81	adenosylhomocysteine
ir2:AG3505	+	1343.00	1582.79	9,4e-81	S-adenosylhomocysteine
ir1:A54040	+	1332.00	1569.81	5,0e-80	adenosylhomocysteine
ir1:A26593	+	1289.00	1519.90	3,3e-77	adenosylhomocysteine
ir2:F40763	+	1274.00	1502.12	3,3e-76	adenosylhomocysteine
ir1:A43629	+	1260.00	1485.57	2,7e-75	adenosylhomocysteine
ir1:A43629	+	1242.50	1464.84	9,9e-74	adenosylhomocysteine
ir1:A45369	+	1237.00	1458.24	9,0e-74	adenosylhomocysteine
ir1:A27655	+	1221.50	1440.05	9,4e-73	adenosylhomocysteine
ir2:S50546	+	1217.50	1434.90	1,7e-72	adenosylhomocysteine
ir2:T32918	+	1203.50	1418.59	1,4e-71	adenosylhomocysteine
ir2:T23833	+	1061.00	1249.25	3,7e-62	adenosylhomocysteine
ir2:H83391	+	1032.00	1235.70	1,8e-61	S-adenosyl-L-homocysteine
ir1:S01302	+	915.50	1212.61	3,1e-60	adenosylhomocysteine
ir2:F08681	+	905.50	1077.60	1,4e-52	adenosylhomocysteine
ir2:C90224	+	840.50	989.80	5,5e-52	adenosylhomocysteine
ir2:S58193	+	797.00	937.87	1,2e-47	probable adenosylhomocysteine
ir2:B72649	+	794.00	934.39	1,4e-44	adenosylhomocysteine
ir2:F7E0401	+	785.50	930.10	4,5e-44	adenosylhomocysteine
ir2:F55518	+	785.50	930.10	4,5e-44	adenosylhomocysteine
ir2:AC1983	+	781.00	918.77	1,0e-43	adenosylhomocysteine
ir2:H69499	+	777.50	914.63	1,0e-43	adenosylhomocysteine
ir2:T771167	+	761.00	895.56	2,8e-42	adenosylhomocysteine
ir2:G69199	+	736.50	866.10	8,8e-41	adenosylhomocysteine
ir2:B75064	+	735.50	864.99	1,0e-40	probable S-adenosyl-L-homocysteine
ir2:F69085	+	734.50	863.92	1,2e-40	adenosylhomocysteine
ir2:B72409	+	696.50	819.24	3,8e-38	adenosylhomocysteine
ir2:G84375	+	590.00	692.66	4,0e-31	adenosylhomocysteine
ir2:S22958	+	372.00	446.76	7,1e-17	adenosylhomocysteine
ir2:F69360	+	239.50	280.38	4,9e-08	adenosylhomocysteine

136 aAlaGluGlnMetLeuThrTrpProAspProAspLysProAlaAsnMetI 153
431 TTGTCGATGATGGTGGTATGCTACACCTCCTCATCTCCAAAGGCTCGAA 480
153 leLeuAspGlyGlyAspAlaThrMetLeuValLeuArgGlyMetGln 169
481 TTCGAACAGCGGTCGCTTCAGAGGCAACAGAGAGCTGACAACTCGA 530
170 TyrGluLysAlaGlyValValProAlaGluLysAspProAlaG 186
531 ATACGCGTGGTCTTCTGCTACATCAAGCAGCTCTCAACCAAGACAAGA 580
186 utrLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCACTGGCACACAGCTGCTGGCGGATGAAGCGTGTTCGGAAGAGACA 630
203 spLysTrpThrLysIleAlaGluSerValLysGlyValThrGluGluThr 219
631 ACAACAGGTGTCACCGCTCTACAGCTCGAGAGGAGGCAACTCCT 680
220 ThrThrGlyValLeuArgLeuThrGlnPheAlaAlaGlyAspLeuAl 236
681 CTTCCAGCCATCAACGCTCAACGCTGTTCACAAAGTCCAAAGTTCGATA 730
236 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGCTGCTCCGCTCTCTATCGATGATGATCAACCGTCTCC 780
253 snLysTrpGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATCGGCGCAAGACAGCTGCTGCTATGGGTACGGCGATGT 830
270 AspAlaLeuIleGlyGlyLysValLeuIleCysGlyTyrglyAspVa 286
831 CGCAGAGGCTGCTCAATCCCTCCCTGCTGGCCAGGCGCTCGCTTATCA 880
286 lGlyLysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTCGACCAATCTGCTGCTCCAGGCTGCCATGAAGGCTAC 930
303 alThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGTCCGCGCTACAGGAGTCTGCTCAAGGATGCTGATATCTTCTGTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValValTh 336
981 ATGCACAGAACTCGCATCATCTCTGTTGACATGATGATGCCAGATGA 1030
336 rAlaThrGlyAsnLysAspIleIleMetLeuGluHisIleLysAlaMetL 353
1031 AGGATAAGGCTATTGTCGGTAACATCGCCCACTTCGATAACGAATGTAT 1080
353 ysAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGGCTCATGAATATCCAGGATCATCAGACATCCCACTCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysPr 385
1131 AGAATACACATGTCGGAGATCCCAAGAT...GCCACAGCTATCTCTCTTC 1177
385 cGlnValAspLeuTrpThrPheGlyAspThrGlyArgSerIleIleVal 402
1178 TTGCTGAGGCGGCTCTTCACTTACCTTGGCTGGCTACAGGTCAACCATCT 1227
402 euSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTCGTTATGATGTCATTACAAACACAGACACTCGCTCAGCTCGACCT 1277
419 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTACGAAAGAGAGGAAATCTCGAAGAGAGGTTTACACACTTCCGAGC 1327
435 utrTrpThrLysAsnAspGluTrpAspAsnGluValTyrglyLeuProLysH 452

1328 ATCTCGATGAAGAGTGGCTCGCTCCACCTCGATCTCTCGATGTCAC 1377
452 isLeuaspGluLysValAlaArgIleHisValGluAlaLeuGlyGlyHis 468
1378 CTTACAAAGCTTACACAGAGCAGGCTGACTACATCAACAGTTCAGTTGA 1427
469 LeuThrLysLeuThrLysGluGlnAlaGluTyrglyLeuGlyValAspValGI 485
1428 GGTCTCTACAGTCTGATGCTTACCGTTAT 1458
485 uGlyProTyrglyLysProAspHisTyrglyTyr 495
seq_name: p12: D87005
seq_documentation_block:
probable S-adenosyl-L-homocysteine hydrolase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: D87005
R: Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holro
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: D87005
A:Molecule type: DNA
A:Residues: 1-492 <STO>
A:Cross-references: GB:AL450380; NID:gl3092881; PIDN:CAC30280.1; GSPDB:GN00147
C:Genetics:
A:Gene: sahH
C:Superfamily: adenosylhomocysteinase
alignment_scores:
Quality: 1537.50 Length: 477
Ratio: 3.932 Gaps: 3
Percent Similarity: 81.971 Percent Identity: 62.264
alignment_block:
US-09-759-990-1 x D87005 ..
Align seg 1/1 to: D87005 from: 1 to: 492
34 GAGTACAGAAATTCGCGACATCAACCTCCATGTTCTCGCGCGTAAGGAAC 83
17 AspPheLysValAlaAspLeuSerLeuAlaAsnPheGlyArgLysGluLe 33
84 TACCTTGTCTGAGAAGAAATGCCAGGTCTTATGTTCTTCTGCTGAGCGTT 133
33 uAspLeuAlaGluTyrgluMetProGlyLeuMetSerLeuArgHisGluT 50
134 ATTCGCGTCTTAAGCCATTTGAAGGTGTCAAGATCTCTGTTCTCCCTCCAC 183
50 yAlaGluValGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 66
184 ATGACACTCCAGACAGCGCTCTCATCGAGACACTCAAGCTCTTGGTGC 233
67 MetThrValGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 83
234 TGATGTCAGATGGCTTCTCTGCAACATCTTCTTACAAAGATACAGCG 283
83 aGluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAlaA 100
284 CTGCTGCTATGTTGTCGCCCAACACACACAGAGAGGAGCCGCGGT 333
100 laAlaAlaValValGlyProTyrglyThrProGluGluProLysGly 116
334 ATCCAGTCTTCTGCTGGAAGGCGCAACACTCCAGAAATATCTGGAGAA 383
:::|||||

```

117 ValProValPheAlaTrpLysGlyGluThrLeuGluGluTyrTrpTrpAl 133
384 CACATACCGCGCTCTCACATGCGCCAGATGGTCTCAAGGCCCA...CAGCAGG 430
133 aAlaGluGlnMetLeuThrTrpProAspLysProValAsnMetI 150
431 TTGTCGATGATGGTGGTGTATCTACACTCTCTCAATCTCCAGGCGCTTCGAA 480
150 leuAspAspGlyGlyAspAlaThrMetLeuValLeuArgGlyValGln 166
481 TTCGAACACGCGGTGCTGTTCCAGAGCCCAACAGAAGCTCAACACCTCGA 530
167 TyrGluLysAlaGlyValValProProAlaGluValAspSerAlaGln 183
531 ATACCGCTCGCTTCTTGCTACACTCAAGCAGGTCTTCAACCAAGACAAGA 580
183 utrPluValPheLeuAsnLeuLeuArgLysArgPheGluThrAspLys 200
581 ACCACTGGCACACAGTGTGCTGCGGCATGAACGGTGTTCCTCCGAAGACA 630
200 lYlYsTrpThrLysIleAlaLysSerValLysGlyValThrGluGluThr 216
631 ACAACAGGTGTCACCGCTCTTACCAGTCTCGAGAAGGAGGCGCAAACTCCT 680
217 ThrThrGlyValLeuArgLeuTyrGlnPheAlaAlaAlaGlyAspLeuAl 233
681 CTTCGCCAGCCATCAACGTCACGACGCTGTTCACAAGTCCAGTTCGATA 730
233 aPheProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 250
731 ACATCTACGGTGGCGGCACATCCCTTATCGATGGTATCAACCGTGCTCC 780
250 snlYsTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 266
781 GATCTCATGATCGCGGCAAGACAGCTCTCGTCATGGGTACGGCGATGT 830
267 AspSerLeuIleGlyGlyLysAsnValLeuIleLysGlyTyrGlyAspVa 283
831 CGGCAAGGGCTCGCTCAATCCCTCCGCGGCAAGCGCTCGCTTATCA 880
283 lGlyLysGlyCysAlaGluAlaAlaLysGlyGlnGlyAlaArgValThrI 300
881 TCACAGACTCGACCCCAATCTCGCTCTCCAGCTGCCATGGAAGCTAC 930
300 leThrGluIleAspProIleAsnAlaLeuGlnAlaLeuMetGluGlyPhe 316
931 CAGTCCGCGCATCGAGGAGTCTGTCAGAGTATGCGATATCTTCGTTAC 980
317 AspValLysArgValGluAspValIleAlaAspSerAspIleValValTh 333
981 ATGCACAGGAACATCGCATATCATCTCTGTTCATGATGATGCCAGATGA 1030
333 rAlaThrGlyAsnLysAspIleIleLeuLeuGluHisMetLysAlaMetL 350
1031 AGGATAAGCTATTCGCTTAACATCGCGCACTTCGATAACGAATTCAT 1080
350 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 366
1081 ACAGATGGCTCATGAATACCCAGGATCAAGCACATCCCAATCAAGCC 1130
367 MetAlaAlaLeuGluArg...SerGlyAlaThrArgLeuAsnIleLysPr 382
1131 AGAATACGACATGTGGGAATTCACAGAT...GGCCAGCTATTCCTCTTC 1177
382 oGlnValAspLeuTrpThrPheGlyAspSerGlyLysSerIleIleValL 399
1178 TTGCTGAGGCGGCTCTTCACTTGGCTGGCTGCTACAGGTCAACCCATCT 1227
399 euserGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 415
1228 TTCGTTATGTCATGTCATTCACAAACACACACTCGCTCAGCTCGACCT 1277
416 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 432

```

```

1278 CTACGAAAGAGAGGAATCTCGAGAGAGAAGTTTACACACTTCGGAAGC 1327
432 utrThrLysAsnAspAspTyrAspAsnGluValTyrArgLeuProLysH 449
1328 ATCTCGATGAAGAGTGCCTCGCCTCCACCTCGGATCTCTCGATCTCCAC 1377
449 lseuAspGluLysValAlaArgValHisValGluAlaLeuGlyGlyGln 465
1378 CTTACAAAGCTTACACAGAAGCGGTGACTACATCAACGTTCCAGTTGA 1427
466 LeuThrLysLeuThrLysAspGlnAlaGluTyrLeuGlyValAspValAs 482
1428 GGTCTCTTCAACAGTCTGATCTACCTACCTTAT 1458
482 pGlyProPheLysProAspHisTyrArgTyr 492

seq_name: p1r2:C71400

seq_documentation_block:
adenosylhomocysteinase (EC 3.3.1.1) [similarity] - Arabidopsis thaliana
N:Alternate names: protein DL3010W; S-adenosyl-L-homocysteine hydrolase
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 02-Sep-2000
C:Accession: C71400; T51608
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Kempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113
A:Accession: C71400
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-485 <BEV>
A:Cross-references: GB:297335; NID:g2244747; PIDN:CAB10173.1; PID:g2244750
R:Belbahri, L.; Elleuch, H.; Villaroel, R.; Inze, D.; Thomas, D.; Thomasset, B.
Plant Physiol. 121, 313, 1999
A:Title: The isolation of an Arabidopsis thaliana cDNA clone encoding S-adenosyl-L-ho
A:Reference number: Z25409
A:Accession: T51608
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-485 <BEL>
A:Cross-references: EMBL:AF059581; PIDN:AAC14714.1
A:Experimental source: cultivar Landsberg erecta
C:Genetics:
A:Gene: SAHH
A:Map position: 4COP9-4G3845
A:Note: DL3010W
C:Superfamily: adenosylhomocysteinase
C:Keywords: NAD; thioether hydrolase
F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F:86,120/Active site: Cys #status predicted

```

alignment_scores:

```

Quality: 1537.00 Length: 487
Ratio: 3.901 Gaps: 4
Percent similarity: 80.903 Percent identity: 62.012

alignment_block:
US-09-759-990-1 x C71400 ..
Align seg 1/1 to: C71400 from: 1 to: 485

```

```

19 ACTGGTGTCCATTCGAGTACAGATTCGCCGACATCAACCTCCATGTTCT 68
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24

```

```

69 CGCCGTAAGAACTTACCTTGTGAGAGGAATGCCAGGCTCTATGG 118
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMetA 41
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 TTCTTGGTGAAGCGTTATTCGCTTCTAAGCCATTGAAGGGTGTCAAGATC 168
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TCCTGTTCCCTCCACATGACAGTCCAGACACCGCTCCTCATCGACACT 218
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 CACAGCTCTGGTCTGATGTCAGATGGCTTCCGCAACATCTTCTCTA 268
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CACAAGATACAGCCGCTGCTGCTATCGTGTGCGCCCAACAGGCACACCA 318
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
91 hrGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
319 GAGAAGCCAGCGGTATCCCAAGTCTCCGCTGGAAGGGCGAACAACACTCCC 368
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
106 .....ValPheAlaTrpLysGlyGluThrLeuG 115
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 AGAATCTGGGAAACACATACCGCGCTCTCACATGGCCAGATGGTCAAG 418
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 nGluYrTrpPcysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 132
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 GCCACAGCAGGTTTCGATGATGTTGGTGATGCTACACTCTCTCATCTCC 468
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
132 lyProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 148
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
469 AAGGCTTCGAA.....TTCGAAACAGCGCGTCTTCACGA 506
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GluGlyValLysAlaGluLeuPheGluLysThrGlyGlnValProAs 165
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
507 GCCAACAAGCTGACAACCTCGAATACCGCTGCTGCTTCTGTGCTACACTCA 556
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleLe 182
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 AGCAGGCTTCAACCAAGACAGAACCACTGGCACACAGTTGTCGCGGC 606
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 ysGluGlyLeuGlnValAspProLysLysThrHisLysMetLysGluArg 198
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
607 ATGAACGCTGTTTCCGAGACACAACAGGTGTCACCGCCCTCTACCA 656
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuYrGL 215
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
657 GCTCAGAGAGGAGGCAAACTCTCTTCCAGCCATCAAGTCAACGACG 706
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAspS 232
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
707 CTGTACAAAGTCCAAAGTTCGATTAACATCTACGGCTGCCGCCCTCTCTT 756
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 erValThrLysSerLysPheAspAsnLeuYrGlyCysArgHisSerLeu 248
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
757 ATCGATGGTATCAACCGTCTTCCGATGTCATGATCGCGCGCAAGACAGC 806
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 TCTCGTCACTGGTTCAGGCGATGTCGGCAGGCTGCGCTCAATCCCTCC 856
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetL 282
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
857 GTGGCCAAAGCGCTCGGTATATCATCAGAACTCGACCCCAATCTGGCT 906
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
907 CTCACAGCTGCCATGAAGGCTACAGGTCGCGCCATCGAGGAAGTCTG 956
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 315
:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

seq_name: pir2:T06764
seq_documentation_block:
adenosylhomocysteine (EC 3.3.1.1) - wheat
N:Alternate names: S-adenosyl-L-homocysteine hydrolase
C:Species: Triticum aestivum (common wheat)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T06764
R:Richards, K.D.; Gardner, R.C.
A:Description: The influence of aluminium on histone, heat shock and S-adenosyl-L-hom
A:Residues: 1-485 <RIC>
A:Reference number: 215795
A:Accession: T06764
A:Status: translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Cross-references: EMBL:L11872; NID:g170772; PIDN:AAA34303.1; PID:g170773
C:Genetics:
A:Gene: SH6.2
C:Superfamily: adenosylhomocysteine
C:Keywords: NAD; thioether hydrolase
F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F:86,120/Active site: Cys #status predicted

alignment_scores:
Quality: 1524.00 Length: 487
Ratio: 3.868 Gaps: 4
Percent Similarity: 80.903 Percent Identity: 61.807

```

alignment_block:

US-09-759-990-1 x T06764

Align seg 1/1 to: T06764 from: 1 to: 485

```

19  ACTGTGCTCCATTCCGAGTACACAAATGCGGACATCAACCTCCATGTTCT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8  thrSerSerGlyArgGluThrLysValLysValLysAspLeuPheGlnAlaAspPh 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69  CGGCCGTAAGGAACCTTACCTTCTGCTGAGAAAGAAATGCCAGGCTCTTAGG 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24  eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMeta 41
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119  TTCCTTCGTCAGGCTTATTCGCTTCAAGCCATTGAAGGGTGTGAGATC 168
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41  laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169  TCTGTCTCCCTCCACATGACAGTCCAGACAGCGCTCCTCCTCGAGACT 218
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58  SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219  CACAGCTCTTGCTGTGATGTACAGATGGGCTTCTCTCAACATCTTCTCTA 268
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74  thrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerS 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269  CACAAGATACACCGCTGCTGCTATCGTTGTCGCGCCCAACAGGCACACCA 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91  erGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319  GAGAACCCAGCGGTATCCAGTCTTCGCTGGAAGGGCGAACAACACTCCC 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106  .....ValPheAlaTrpLysGlyGluThrLeuGI 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369  AGAATACCTGGGAGAACATACATCCGCTCTCACATGCGCCAGTGGTCAAG 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115  uGluTrpTrpCysThrGluArgCysLeuAspTrpGlyValGlyGly 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419  GCGCCACAGCGTTCGATGATGTTGGTGTGATGCTACATCTTCATCTCC 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132  lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 148
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469  AAGGCTTC.....GAATTCGAAACAGCGGCTGTCTCCAGA 506
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
149  GluGlyValLysAlaGluGluPheGluLysSerGlyLysValProAs 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507  CCAACAGACGCTGACACCTCGAATACCGCTGCGTTCCTTGTCTACACTCA 556
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165  pProGluSerThrAspAsnProGluPheLysIleValLeuThrIleIleA 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
557  AGCAGGCTCTCAACCAACAGACAGAACCTGCGCACACAGTTGTCGCGGC 606
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182  rGAspGlyLeuLysThrAspAlaSerLysTrpArgLysMetLysGluArg 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
607  ATGAACGCTGTTTCGAGAGACACACAGAGTGTCCACGCGCTCTACCA 656
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
199  LeuValGlyValSerGluLysThrThrGlyValLysArgLeuTrpGI 215
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
657  GCTCGAGNAGAGGCAACCTCTCTCCAGCCATCAACGTCACAGCAGC 706
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215  nMetGlnGluSerGlyThrLeuLeuPheProAlaIleAsnValAsnAsps 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
707  CTGTTTACAAGTCAAGTTCGATACATCTACGCTGCGCGCCACTCCCTT 756
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232  erValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
757  ATCGATGGTATCAACCGGCTTCGATGTCATGATCGCGCGCAAGACAGC 806
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249  ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
807  TCTCGTCAATGGTTACGCGGATGTCGCGCAAGCGCTGCGCTCAATCCCTCC 856
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265  aValValCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaLeuL 282

```

```

857  GTGCCCAAGGCGCTGCTTATCATCACAGAACTCGACCCCAATCTCGGCT 906
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282  ysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
907  CTCAGGCTGCCATCGAAGGCTACAGGTCGCCGCGCATCGAGGAAGTCGT 956
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299  LeuGlnAlaLeuMetGluGlyIleGlnIleLeuThrLysAspValVa 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
957  CAAGGATGTCATATCTTCGTTACATGCACAGGAACACTCGGATATCATCT 1006
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
315  lSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 332
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1007  CTGTTGACATGATGCCGCCAGATGAAGGATAAGGCTATTCTCGGTAACATC 1056
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332  etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1057  GGCACCTTCGATACAGAAATGATACAGATGGCTCTCATGAAATACCCAGG 1106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349  GlyHisPheAspAsnGluIleAspMetAsnGlyLeuGluThrTyProGI 365
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1107  CATCAAGCACATCCCAATCAAGCCAGAAATACACATGTGGGAATCCCAAG 1156
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365  yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProG 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1157  ATGCCAC...GCTATCTCTCTTGTGCTGAGGCGGCTTCTTAACCTT 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
382  luthrLysThrGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1204  GGTGCTGCTACAGTCAACCATCTTTCGTTATGTCAATGTCTATTCACAAA 1253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399  GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1254  CCAGACACTCGCTCAGCTCAGCTCTAGAAAAGAGA.....CGAAATC 1297
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415  nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysAlaSerGlyLysT 432
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1298  TCAGAGAAGAGGTTTACACACTTCCGAGCATCTCGATGAAGAAGTCGCT 1347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432  YrGluLysLysValTyValLeuProLysHisLeuAspGluLysValAla 448
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1348  CGCTCCACCTCGGATCTCGATGTCCTCCACCTTACAAAGCTTACACAGAA 1397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449  AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysSe 465
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1398  GCAGGCTGCTACATCAACGTTCCAGTTGAGGTCCTTACAGTCTGATG 1447
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465  rGlnSerAspTyIleSerIleProIleGluGlyProTyLysLeuArgL 482
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448  CTTACCGTTAT 1458
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482  euTyArgTyArg 485

```

seq_name: pir2:S38379

seq_documentation_block:

adenosylhomocysteinase (EC 3.3.1.1) - Madagascar periwinkle

N/Alternate names: S-adenosyl-L-homocysteinase hydrolase

C/Species: Catharanthus roseus (Madagascar periwinkle)

C/Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C/Accession: S38379

R/Schroeder, G.; Waitz, A.; Hotze, M.; Schroeder, J.

submitted to the EMBL Data Library, October 1993

A/Description: cDNA for stress-induced S-adenosyl-L-homocysteinase hydrolase from Madag

A/Reference number: S38379

A/Accession: S38379

A/Molecule type: DNA

A/Residues: 1485 <SCH>

A/Cross-references: EMBL:Z26881; NID:g407411; PIDN:CAA81527.1; PID:g407412

C/Genetics:

A/Gene: SAHH

C/Superfamily: adenosylhomocysteinase

C/Keywords: NAD; thioether hydrolase

F: 263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
F: 86,120/Active site: Cys #status predicted

alignment_scores:
Quality: 1520.00 Length: 488
Ratio: 3.848 Gaps: 6
Percent Similarity: 80.943 Percent Identity: 62.295

alignment_block:
US-09-759-990-1 x S38379 ..

Align seg 1/1 to: S38379 from: 1 to: 485

```
19 ACTGGTGTCCATTCAGTACAGATTGCCGACATCAACCTCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCGTAAGGAACCTTACCCTTGCCTGAGAGGAATGCCAGGTCTTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuGluAlaGluValGluMetProGlyLeuMets 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTCCTGCGAGGCTTATCCGCTTCTTAAGCCATTGAAGGTGTCAATC 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 exCysArgAlaGluPheGlyProSerGlnProPheLysGlyAlaLysIle 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 TCTGGTCTCCCTCCATGACACTCCAGACAGCGCTCTCATCGAGACACT 218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 CACAGCTCTTGTGTGATGTCAGATGGCTTCTGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSer 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 CACAGATACACCGCTGTCTATCTGTTGTCGGCCCAACAGCCACACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hrGlnGluHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GAGAAGCCAGCGGTATCCAGTCTTCGCTGCGCTGGAAGGGGAAACACTCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaTrpLysGlyGluThrLeuGlu 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 AGAATCTGGGAGAACATACCGCTCTCATAGG...CCAGATGGTC 415
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProAspGly. 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 AAGGCCACAGCAGGTGTCGATGGTGGTGCATGCTACACTCCTCATC 465
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
132 ..GlyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIle 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 TCCAGGGGCTC.....GAATTGGAACACAGCGGTGCTGTTC 503
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 HisGluGlyValLysAlaGluGluTyrLysLysAsnGlyAlaLeuPr 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 AGAGCCAAACAGAGCTGACACCTCCGAATACCGCTGGTCTTCTGCTACAC 553
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 aspProSerSerThrAspAsnAlaGluPheGlnIleValLeuThrIle 181
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554 TCAACGAGGTCTTCAACCAAGACAAGAACACCTGGCACACAGTTGCTGCC 603
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 leArgAspGlyLeuLysSerAspProThrLysTyrThrArgMetLysGlu 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
604 GCATGAAGGCTTTCCGAGAGACAACACAGGTGCTCCACCGCTCTA 653
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 ArgLeuValGlyValSerGluGluThrThrGlyValLysArgLeuTy 214
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654 CCAGCTCGAGAGGGGAACTCTCTTCCAGCCCATCAACGTCAACG 703
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 rGlnMetGlnAlaAsnGlyThrLeuLeuPheProAlaIleAsnValAsnA 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704 ACCTGTTACAAAGTCCAGTTCGATAACACTACGGCTCCCGCCACTCC 753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 spSerValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSer 247
```

```
754 CTTATCGATGGTATCAACCGTGTTCGATGTCATGTCGGCGCAAGAC 803
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 LeuProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVa 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
804 AGTCTCGTCATGGTTACGGCGATGTCGGCAAGGGTGCCTCAATCCC 853
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 lAlaValAlaGlyTyrGlyAspValGlyLysGlyCysAlaAlaL 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
854 TCCGTGGCCAAAGCGCTCGGTTATCATCACAGAACTCGACCCATCTGC 903
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
281 euLysGlnAlaGlyAlaArgValIleValThrGluIleAspProIleCys 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
904 GCTCTCCAGGCTGCCATGGAAGGTACCAGTCCGCCGCATCGAGGAAGT 953
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 AlaLeuGlnAlaThrMetGluGlyLeuGlnValLeuThrLeuGluAspVa 314
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954 CGTCAGGATGTCGATATCTTCGTTACATGCACAGAACTCCGATATCA 1003
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 lValSerGluAlaAspIlePheValThrThrGlyAsnLysAspIleI 331
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1004 TCTCTGTTGACATGATGCCCGCAGATGAAGGATAAGCTATTGTCGGTAAC 1053
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
331 leMetValAspHisMetArgLysMetLysAsnAlaIleValCysAsn 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1054 ATCGGCCACTTCGATAAGAAATGATGACAGATGGCCTCATGAATACCC 1103
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348 lIleGlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyPr 364
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1104 AGCATCAAGCACATCCCAATCAAGCCAGATACGACATGTGGGAATTC 1153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
364 oGlyValLysArgIleThrIleLysProGlnThrAspArgTrpValPheP 381
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1154 CAGATGGCCAGCT...ATCCTCTCTTGTGTGAGGCGCGCTTCTTAAC 1200
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
381 roAspThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsn 397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201 CTTGGCTCGCTACAGTACCCATCTTTCGTTATGTCATGTCATTCAC 1250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 LeuGlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheTh 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1251 AAACAGACACTCGCTCAGCTCGACTCTACGAAAGAGA.....GGAA 1294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 rAsnGlnValIleAlaGlnLeuGluLeuTrpAsnGluArgLysThrGlyL 431
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1295 ATCTCGAAGAAGGTTTACACACTTCGAGACATCTCGATGAAGAATC 1344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 ystyrGluLysLysValTyrValLeuProLysHisLeuAspGluLysVal 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1345 GCTCGCTCCACCTCGATCTCGATGTCACCTTACAAAGCTTACACA 1394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
448 AlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrly 464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1395 GAAGCAGGCTGACTACATCAACGTTCCAGTTCAGGGTCTTACAAAGTCG 1444
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
464 saspglnAlaAspTyrIleSerValProIleGluGlyProTyrLysProA 481
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1445 ATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 laHisTyrArgTyr 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
seq_name: pir2:S71621
```

seq_documentation_block:

adenosylhomocysteinase (EC 3.3.1.1) - Phalaenopsis sp.

N;Alternate names: S-adenosyl-L-homocysteine hydrolase

C;Species: Phalaenopsis sp.

C;Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 22-Jun-1999

C;Accession: S71621

R;Preisig-Mueller, R.; Gnaul, P.; Kindl, H.

Arch. Biochem. Biophys. 317, 201-207, 1995

A;Title: The inducible 9,10-dihydrophenanthrene pathway: characterization and express

A;Reference number: S71619; MUID:95177653

A:Accession: S71621
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-485 <PRE>
 A:Cross-references: EMBL:X79905; NID:G758246; PIDN:CAA56278.1; PID:G758247
 A:Experimental source: clone SH511
 C:Superfamily: adenosylhomocysteinase
 C:Keywords: NAD; thioether hydrolase
 F:263-292/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:86,120/Active site: Cys #status predicted

alignment_scores:
 Quality: 1509.00 Length: 482
 Ratio: 3.889 Gaps: 4
 Percent Similarity: 80.498 Percent Identity: 62.656
 alignment_block:
 US-09-759-990-1 x S71621 ..
 Align seg 1/1 to: S71621 from: 1 to: 485

```

34 GAGTACAGAAATGCGGACATCAACATCCATGTTCTCGCGCGTAAGGAAC 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 GluTyrLysValLysAspLeuSerGlnAlaAspPheGlyArgLeuGluI 29
84 TACCTTGCTGAGAAAGGAATGCCAGGTCTTATGTTCTTCCTGAGCGTT 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
29 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgAlaGlu 46
134 ATTCGGCTTCTAAGCCATTGAAGGTGTGAGAATCTCTGGTTCCTCCAC 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 heGlyProSerGlnProPheLysGlyAlaArgIleSerGlySerLeuHis 62
184 ATGACATGTCAGACAGCGCTCCTCATCGAGACATCAAGCTCTTGGTGC 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 79
234 TGATGCAGATGGCTTCCTGCAACATCTTCTTACACAAGATACAGCG 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 aGluValArgTyrCysSerCysAsnIlePheSerThrGlnAspHisAla 96
284 CTGCTCTATCTGTTGCGGCCACACAGGACACACAGAGAACCCAGCG 333
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 laAlaIleAlaIleAlaArgAspSerAlaAla..... 105
334 ATCCAGATCTTCGCTGGAGGCGGAACACTCCAGATACTGGGAGAA 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
106 .....ValPheAlaIlePheGlyGluThrLeuGlnGluTyrTrpCy 120
384 CACATACCGCTCTCACATGCGCAGATGTCAGGCCACACAGAGTTG 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 sThrGluArgCysLeuGluTrpGlyAlaGlyGlyGlyProAspLeuIle 137
434 TCGATGATGGTGGTGTATGATGATGATGATGATGATGATGATGATG 477
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 alAspAspGlyGlyAspAlaThrLeuLeuIleHisGluGlyValLysAla 153
478 .....GAATTCGAACAGCGGTGCTGTTCCAGAGCCACAGAGCTGA 521
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 GluGluGluTyrGluLysAsnGlyLysIleProAspProAlaSerThr 170
522 CAACCTCGATACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 pAsnAlaGluPheGlnIleValLeuGlyLeuIleArgAspSerLeuSer 187
572 AAGACAGAACCCACTGGCACAGATGCTGCTGCGGCATGACGGTGTTC 621
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 alAspProLysLysTyrArgArgMetLysGluArgLeuValGlyValSer 203
622 GAAGAGACAAACAGGTGTCCACCGCTCTACACAGCTGAGAGGAGGG 671
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 GluGluThrThrThrGlyValLysArgLeuTyrGlnMetGlnTyrSerGI 220

```

```

672 CAACATCTCTTCCACGCCATCAAGCTCAACGACGCTGTTACAAAGTCCA 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 237
722 AGTTCCGATAACATCTACGGCTGCGCGCACCTCCCTTATCGATGGTATCA 771
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 yspPheAspAsnLeuTyrGlyCysArgHisSerLeuProAspGlyLeuMet 253
772 CGTGTTCGATGTCATGATCGGGCGGCAAGACAGCTCTCTGTCATGGGTTA 821
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ArgAlaThrAspValMetIleAlaGlyLysValAlaValValCysGlyTy 270
822 CGGCGATGTCGCAAGGCGTCCGCTCAATCCCTCCGTCGCGCAAGCGGTC 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 rGlyAspValGlyLeuGlyCysAlaAlaLeuLysThrAlaGlyAlaA 287
872 GCGTTATCATCAGAACTCGACCCCAATCTCGCTCTCCAGGCTGCCATG 921
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 rgValIleValThrGluIleAspProIleCysAlaLeuGlnAlaLeuMet 303
922 GAAGCTACCGTCCGCGCATCGAGGAAGTCTCAAGGATGTCGATAT 971
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 GluGlyLeuProValLeuArgLeuGluAspValValSerGluAlaAspI 320
972 CTTCTGTTACATGCACAGGAACCTCGCATATCATCTCTGTTGACATGATG 1021
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ePheValThrThrGlyAsnLysAspIleIleMetValAspHisMetA 337
1022 CCCAGATGAAGGATAAGGCTATTGTCGTAACATCGGCCACTTCGATAAC 1071
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 rgLysMetLysAsnAlaIleValCysAsnIleGlyHisPheAspAsn 353
1072 GAAATGATACAGATGGCTCATGAATACCCAGGATCAAGCACATCCC 1121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GluIleAspMetLeuGlyLeuGluSerPheProGlyValLysArgIleTh 370
1122 AATCAAGCCAGAATACGACATGTGGAAATCCAGATGGCCAGCT....A 1168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 rIleLysProGlnThrAspArgValPheProAspThrAsnSerGlyI 387
1169 TCCCTCTCTTCTGAGGCGCCCTTAACTTGGCTGGCTGACAGGT 1218
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 leuValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 403
1219 CACCATCTTCTGTTATGTCATGTCATTACAAACCCAGACACTCGCTCA 1268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 HisProSerPheValMetSerSerSerPheThrAsnGlnValIleAlaGI 420
1269 GCTCGACCTCTACGAAAGAGA.....GGAAATCTCGAGAAAGGTTT 1312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 nLeuGluLeuTrpLysGluArgAlaSerGlyLysTyrGluLysLysVal 437
1313 ACACACTTCGGAACATCTCCATGAAGAAGTCGCTGCGCTCCACCTCGGA 1362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
437 yValLeuProLysHisLeuAspGluLysValAlaAlaLeuHisLeuGly 453
1363 TCCTCGATGTCACCTTACAAAGCTTACAGAGAGCGGCTGACTACAT 1412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 LysLeuGlyAlaLysLeuThrLysLeuThrProSerGlnAlaAspTyrI 470
1413 CAACGTTCCAGTTGAGGTCCTTCAACAGCTGATGCTTACCGTTAT 1458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 eSerValProValGluGlyProTyrLysProAlaHisTyrArgTyr 485

```

seq_name: pir2:D82730

seq_documentation_block:

adenosylhomocysteinase Xf1037 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: D82730
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: D82730

A:Status: preliminary

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-446 <SIM>

A:Cross-references: GB:AE003941; GB:AE003949; NID:g9105978; PIDN:AAF83847.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1037

C:Superfamily: adenosylhomocysteinase

alignment_scores:

Quality: 1406.00 Length: 457

Ratio: 3.790 Gaps: 5

Percent Similarity: 81.182 Percent Identity: 60.613

alignment_block:

US-09-759-990-1 x D82730 ..

Align seg 1/1 to: D82730 from: 1 to: 446

103 ATCCAGGCTTATGGTTCTTCGTGAGGCTTATTCGGCTTCTAAGCCATT 152

1 MetProglyLeuMetSerIleArgArgIlySerAlaSerLysGlnProLe 17

153 GAAGGTTGTCAGATCTCTGGTTCCCTCCACATGACAGTCCAGACAGCCG 202

17 ulysGlyValArgValThrGlySerLeuHisMetThrIleGlnThrAlay 34

203 TCCTCATCGAGACTCACACCTCTTGTGGTGTGATGTCAGATGGGCTTC 252

34 alLeuIleGluThrLeuLysAspIleGlyAlaAspValArgTrpAlaSer 50

253 TGCACATCTTCTTACACAGATACAGCCGCTGCTGCTATCGTTGTCGG 302

51 CysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAlaThr.. 66

303 CCCAACAGGCACACAGAGAGCCAGCCGGTATCCAGTCTTCGCTGGA 352

67SerglyThrProValPheAlaTrpL 75

353 AGGCGGAACACTCCCAAGTAATCTGGGGAACACATACCGGCTCTCAC 402

75 ysGlyGluThrLeuGluGluThrTrpAspCysThrLeuGlnAlaLeuThr 91

403 TGG.....CCAGATGGT.....CAAGGCCACACAGAGTGTTCGATGA 440

92 PheThrLeuSerAspGlyThrLeuThrGlyProGluLeuIleValAsp 108

441 TGGTGGTGTGCTACACTCTCATCTCCAAAGGGCTTCGAATTCGAAACAG 490

108 pGlyGlyAspAlaThrLeuLeuIleHisysGlyTyrGluLeuLuanG 125

491 CCGGTGCT...GTTCCAGAGCCACAGAGAGCTGACAACTCGAATACCGC 537

1438 AAGTCTGATGCTTACGGTTAT 1458
 ||| ::: |||||
 440 LysProGluHisTyrArgTyr 446

seq_name: pir2:D97362

seq_documentation_block:

adenosylhomocysteine (S-adenosyl-L-homocysteine hydrolase) (adocycase) [imported] - Ag
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: D97362

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:11743194

A:Accession: D97362

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-466 <KOR>

A:Cross-references: GB:AE007869; PIDN:AAK85853.1; PID:gi1514892; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_46

A:Map position: circular chromosome

C:Superfamily: adenosylhomocysteine

alignment_scores:

Quality: 1382.00 Length: 478
 Ratio: 3.627 Gaps: 6
 Percent Similarity: 79.707 Percent Identity: 58.577

alignment_block:

US-09-759-990-1 x D97362 ..

Align seg 1/1 to: D97362 from: 1 to: 466

34 GAGTACAGAAATGCGGACATCAACCTCCATGTCCTCGGCGGTAAAGAACT 83

6 AspTyrIleValAlaAspIleAsnLeuAlaPheGlyArgLysGluLe 22

84 TACCTCTGCTGAGAGGAAATGCCAGGTCTTATGGTCTCTCGTGAGCGTT 133

22 uAspIleAlaGluThrGluMetProGlyLeuMetSerCysArgLysGluP 39

134 ATTCCGCTTAAAGCATTAAGGGTGTGAGAAATCTCTGGTCCCTCCAC 183

39 heGlyGluSerLysProLeuLysGlyAlaArgIleThrGlySerLeuHis 55

184 ATGACAGTCCAGACAGCGCTCTCATCGAGACACACACAGCTCTGGTGC 233

56 MetThrIleGlnThrGlyValLeuIleGluThrLeuLysGluLeuGlyAl 72

234 TGATGTCAGATGGCTTCTGCAACATCTCTCTACAAAGATACAGCGG 283

72 aGluIleArgTyrAlaSerCysAsnIlePheSerThrGlnAspHisAla 89

284 CTGCTGCTATCGTTGTGCGGCCACACAGGCACACAGAGCCAGCCGGT 333

89 laAlaIleAlaAla.....AlaGly 96

334 ATCCAGCTCTTGGCTGGAAGGGGAAACACTCCAGCAATCTACTGGAGAA 383

97 IleProValPheAlaValLysGlyGluSerLeuThrGluTyrTrpGluTyr 113

384 CACATACCGGCTCTCATATGCCAGATGGTCAAGGCCACACAGAGGTG 433

113 rThrAspLysIlePheGlnThrTrpAspGlyLeuSerAsnMetIleL 130

434 TCGATGATGGTGTGATCTACATCTCTCATCTCCAGGGCTTCGAAATC 483

130 euAspAspGlyGlyAspAlaThrMetTyrIleLeuLeuGlyAlaArgAla 146

484 GAAACAGCC...GGTGTCTGTTCCAGAGCCA...ACAGAAGCTGACAACCT 527

147 GluAlaGlyGluaspValLeuSerAsnProGlySerGluGluGluGlu 163

528 CGAATACCCGCTGGCTTCTTGTACACATCAAGCAGGTCTTCAACCAAGACA 577

163 eLeuPheAlaGlnIleAsnLysArgLeuLysAla.....S 175

578 AGAACCACTGGCACACAGTGTCTGCC...GGCATGAACGGTGTTCCTCGAA 624

175 erProGlyTyrPheThrLysGlnArgAspAlaLeuLysGlyValThrGlu 191

625 GAGACAACAACAGGTGTCCACCGCTCTACCAAGTCTCGAGAGGAGGCGAA 674

192 GluThrThrThrGlyValHisArgLeuTyrAspLeuAlaLysGlyLe 208

675 ACTCTCTTCCAGCCCATCAACGCTCAACGAGCTGTGTACAAAGTCCAGT 724

208 uLeuProPheProAlaIleAsnValAsnAspSerValThrLysSerLysP 225

725 TCGATAACATCTACGGCTGCGGCCACTCCCTTATCGATGGTATCAACGT 774

225 heAspAsnLysTyrGlyCysLysGluSerLeuValAspGlyIleArg 241

775 GCTTCCGATGTCTATGATCGCGCGCAAGACAGCTCTCGTCATGGGTACGG 824

242 AlaThrAspValMetMetAlaGlyLysValAlaValValCysGlyTyrG 258

825 CGATGTGCGCAAGGGTGGCTCATCTCCCTCGGCGCAAGGCGCTCGCG 874

258 yAspValGlyLysGlySerAlaAlaSerLeuGlnGlyAlaGlyAlaArg 275

875 TTATCATCACAGAACTCGACCAATCTGCGCTCTCCAGGCTGCCATGAA 924

275 alLysValThrGluIleAspProIleCysAlaLeuGlnAlaAlaMetAsp 291

925 GGCTACCAAGTCCGCGCATCGAGGAAGTGTCTCAAGGATGTCTGATATCTT 974

292 GlyPheGluValValArgLeuGluAspValValSerSerAlaAspIlePh 308

975 CTTTACATGCACAGGAAACTGCCATATCTCTGTGTGACATGATGGCCC 1024

308 eIleThrThrThrGlyAsnLysAspValIleArgIleGluHisMetArg 325

1025 AGATGAAGGATAAGGCTATTGTGCGTAACATCGGCCACTTCGATTAACGAA 1074

325 luMetLysAspMetAlaIleValGlyAsnIleGlyHisPheAspAsnGlu 341

1075 ATTGATACAGATGGCTCATGAAATACCCAGGCATCAAGCACATCCCAAT 1124

342 IleGlnValAlaSerLeuArg.....AsnLeuLysTrpThrAsnI 355

1125 CAAGCCAGAAATACGATGTGGGAATCCAGATGGCCACGCTATCTCTCC 1174

355 elLysProGlnValAspMetIleGluPheProLysGlyAsnArgIleIleL 372

1175 TTTCTGCTGAGGCGCGCTTCTTAAGCTTGGCTGCGCTACAGGTCACCCA 1224

372 euLeuSerGluGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisPro 388

1225 TCTTTGTTTATGTCATGTCATTCACAAACGACACTCGCTCAGCTCGA 1274

389 SerPheValMetSerAlaSerPheThrAsnGlnValLeuGlyGlnIleG 405

1275 COTCTACGAAAGAGAGGAAATCTCCAGAGAGAGGTTTACACACTTCGCA 1324

405 uLeuPheThrLysProGlyGluTyrLysAsnGluValTyrValLeuProL 422

1325 AGCATCTCGATGAAGAAGTCGCTCGCTCCACCTCGGATCTCTCGATGTC 1374

422 yHisLeuAspGluLysValAlaArgLeuHisLeuGluLysLeuGlyVal 438

389 MetSerAlaSerPheThrAsnGlnThrLeuAlaGlnIleGluLeuTprTh 405
1284 AAAGAGAGAAATCTCCAGAGAGAGAGTTTACACACTCCGAGGATCTCG 1333
405 rAsnLysAlaLysTyrGluAnGlnValTyrThrLeuProLysHisLeuA 422
1334 ATGAAGAGTGCCTCCACCTCCGATCTCTCGATGTCACCTTACA 1383
422 spGluLysValAlaPheLeuHisLeuGluLysLeuGlyAlaLysLeuThr 438
1384 AGCTTACACAGAGAGGCTGATCTACATCAACAGTTCAGTTGAGGTC 1433
439 ThrLeuArgLysAspGlnAlaAspTyrIleGlyValProGluAlaGlyPr 455
1434 TTACAAGTCTCATCTACCGCTTAT 1458
455 oPheLysProAspHisTyrArgTyr 463
seq_name: pir1:A46035
seq_documentation_block:
adenosylhomocysteine (EC 3.3.1.1) - Rhodobacter capsulatus
N; Alternate names: S-adenosyl-L-homocysteine hydrolase
C; Species: Rhodobacter capsulatus
C; Date: 04-Mar-1994 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
C; Accession: A46035; A36863
R; Sganga, M.W.; Aksamit, R.R.; Cantoni, G.L.; Bauer, C.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 6328-6332, 1992
A; Title: Mutational and nucleotide sequence analysis of S-adenosyl-L-homocysteine hydrolase
A; Reference number: A46035; MUID:92335291
A; Accession: A46035
A; Molecule type: DNA
A; Residues: 1-462 <SQA>
A; Note: sequence extracted from NCBI backbone (NCBIN:108691, NCBI:P:108696)
R; Bugg, J.J.; Sganga, M.W.; Bauer, C.E.
J. Bacteriol. 176, 61-69, 1994
A; Title: Nucleotide sequence and characterization of the Rhodobacter capsulatus hvrB gene
A; Reference number: A36863; MUID:94110241
A; Accession: A36863
A; Molecule type: DNA
A; Residues: 1-13 <BUG>
A; Cross-references: GB:L23836; NID:g577613; PIDN:AAA53540.1; PID:g577614
C; Genetics:
A; Gene: ahcy
C; Function:
C; Superfamily: adenosylhomocysteinease
C; Keywords: NAD; thioether hydrolase
F; 247-276/Region: beta-alpha-beta NAD nucleotide-binding fold

184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACACTCACAGCTCTTGTGTC 233
53 MetThrIleGlnThrAlaValLeuIleGluThrLeuLysAlaLeuGlyAl 69
234 TGNATGCATGAGGCTCTCCGACACACTCTCTACACAGATACACCGC 283
69 aspValArgTyrPheSerCysAsnIlePheSerThrGlnAspHisAla 86
284 CTGCTCTATCTGCTGCGCCCAACAGGCACACAGAGAGAGCCAGCCGCT 333
86 laAlaAlaIleAlaAlaGly.....Gly 93
334 ATCCAGTCTTCCGCTGGAGAGGGAACACTCCACAAATACTGGAGAA 383
94 ThrProValPheAlaValLysGlyGluThrLeuGluGluTyrTrpAlaTy 110
384 CACATACCGGCTCTCATGTCAGGTCAGATGTCAGAGGCCACAGCAGGTTG 433
110 rThrAspLysIlePheGlnPheProGluGlyThr....CysAsnMetIleL 126
434 TCGATGATGCTGCTGATCTACATCTCTCATCTCCAAAGGCTTCGAATTC 483
126 euAspAspGlyGlyAspAlaThrLeuTyrIleLeuLeuGlyAlaArgVal 142
484 GAAACA.....GCCGCTGCTGTTCCAGAGCCCAACAGAGCTGCAACCT 527
143 GluAlaGlyGluThrAspLeuIleAlaThrProThrSerGluAspGluVa 159
528 CGAATACCGCTGCTTCTGCTACACTCAAGCAGGCTCTTCAACCAAGACA 577
159 l.....Cys.....LeuPheAsnGlnIleL 166
578 AGAACAC.....TGGCACACA....GTTGCTGCGCGC 606
166 YsLysArgMetValGluSerProGlyThrPheThrGlnArgAlaAla 182
607 ATGACAGTCTTCCGAGAGACACACAGCTGTCACCGCTCTTACCA 656
183 IleLysGlyValSerGlyGluThrThrThrGlyValHisArgLeuTyrAs 199
657 GCTCGAGAGAGGAGCAAACTCTCTTCCAGCCATCAACGTCACACGAC 706
199 pLeuHisLysLysGlyLeuLeuProPheProAlaIleAsnValAsnAsps 216
707 CTGTTACAAAGTCCAGTTCGATACATCTACGCTGCGCGCTCCCTT 756
216 erValThrLysSerLysPheAspAsnLysTyrGlyLysGluSerLeu 232
757 ATCGATGATCAACCGCTGCTGCTGATGTCATGTCGCGCGGCAAGACAGC 806
233 ValAspGlyIleArgAlaThrAspValMetMetAlaGlyLysValAl 249
807 TCTCGTCATGGTTCAGCGGATGTCGCAAGGCTGCGCTCAATCCCTCC 856
249 aValValCysGlyTyrGlyAspValGlyLysGlySerAlaAlaSerLeuA 266
857 GTGCCAAGCGCTCGCTTATCATCACAGAACTCGACCCCAATCTCGCT 906
266 rgGlyAlaGlyAlaArgValLysValThrGluValAspProLysAla 282
907 CTCAGGCTGCTGAGAGGCTACAGGTCGCGCTCCGCTCATCGAGAAATCGT 956
283 LeuGlnAlaAlaMetAspGlyPheGluValValLeuGluAspValVa 299
957 CAAGATGCTCATATCTTCTGCTTACATGCACAGGAACTCCGATATCATCT 1006
299 lalaAspAla...IlePheIleThrThrGlyAsnLysAspValIleA 315
1007 CTGTTGACATGATGCTGCCAGCATGAAGGCTATTTCGCGTAACATC 1056
315 rgIleGluHisMetArgGluMetLysAspMetAlaIleValGlyAsnIle 331
1057 GGCCACTTTCGATACGAAATTTGATACAGATGCGCTCATGAATACCCAGG 1106

alignment_scores:
Quality: 1352.50 Length: 484
Ratio: 3.636 Gaps: 9
Percent Similarity: 76.860 Percent Identity: 59.298
alignment_block:
US-09-759-990-1 x A46035 ..
Align seg 1/1 to: A46035 from: 1 to: 462
34 GAGTACAGAAATTCGACATCAACCTCCATGTTCTCGCGCTGAAGNACT 83
3 AspTyrIleValLysAspIleLysLeuAlaGluPheGlyArgLysGluLe 19
84 TACCCTTGTGTGAGAGAAATGCCAGGCTCTTATGTTCTTCGTGAGCGTT 133
19 uAspIleAlaGluThrGluMetProGlyLeuMetAlaCysArgGluGluP 36
134 ATTCCGCTTCTAAGCCATGAAGGCTGTCAGATCTCTGTTCCCTCCAC 183
36 heGlyProSerGlnProLeuLysGlyAlaArgIleAlaGlySerLeuHis 52


```

|||||
332 GlyHisPheAspAsnGluIleGlnValAlaAlaLeuLysAsnHis..... 346
1107 CATCAAGCACATCCCAATCAAGCCAGAAATAGACATGTGGAAATCCCCAG 1156
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
347 ....LysTrpThrAsnIleLysAspGlnValAspMetIleGluMetProS 362
1157 ATGGCCACGCTATCCCTCTCTGCTGAGGCCGCTTCTTAACTTGGC 1206
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
362 erGlyAlaArgIleIleLeuLeuSerGluGlyArgLeuLeuAsnLeuGly 378
1207 TCGCTACAGGTACCCATCTTTCGTTATGTCATGTCTATTCACAAACCA 1256
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 AsnAlaThrGlyHisProSerPheValMetSerAlaSerPheThrAsnG1 395
1257 GACACTCGCTCGAGCTCGACCTCTACGAAAGAGAGAAATCTCGAAGA 1306
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
395 nValLeuAlaGlnIleGluLeuTrpThrLysGlyAlaGluTrpGlnProG 412
1307 AGGTTTACACTTCGAGACATCTCGATGAAGAAGTCGCTGCCCTCCAC 1356
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 lyValTrpIleLeuProLysSerLeuAspGluLysValAlaArgLeuHis 428
1357 CTCGATCTCTGATCTCCACCTTACAAAGCTTACACAGAGCAGCTGA 1406
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 LeuLysLysIleGlyValLysLeuThrThrLeuArgProAspGlnAlaG1 445
1407 CTACATCAACGTTCCAGTTGAGGGTCTCTTACAACTCTGATGCTTACCGTT 1456
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 uTyrIleGlyValThrValGluGlyProPheLysSerAspHisTyrArgT 462
1457 AT 1458
||
462 Yr 462

```

seq_name: pir2:AG3505

seq_documentation_block:

adenosylhomocysteinase (SC 3.3.1.1) [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

R:Accession: AG3505

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, P.; Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AG3505

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-481 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL53210.1; PID:g17984086; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI2029

A:Map position: 1

C:Superfamily: adenosylhomocysteinase

C:Keywords: thioether hydrolase

alignment_scores:

Quality:	1343.00	Length:	475
Ratio:	3.562	Gaps:	4
Percent Similarity:	79.368	Percent Identity:	57.053

alignment_block:

US-09-759-990-1 x AG3505 ..

Align seg 1/1 to: AG3505 from: 1 to: 481

34 GAGTACAGAATTCGCGACATCAACCTCCATGTTCTCGCGGTAAAGAACT 83

:||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

21 AspPheValValLysAspIleSerLeuAlaAspTrpGlyArgGluLeu 37

```

84 TACCCTTGCTGAGAAGAAATGCCAGGTCCTTATGCTTCTCTCGTGAGCGTT 133
|:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
37 uaspIleAlaGluThrGluMetProGlyLeuMetAlaAlaArgGluGluP 54
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
134 ATTCCGCTTCTTAAGCCATTAAGGGTGTGAGAATCTCTGGTTCCCTCCAC 183
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
54 heglyLysSerGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 70
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
184 ATGACAGTCCAGACAGCCGCTCTCATCGAGACATCAGACTCTTGGTGC 233
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 MetThrIleGlnThrAlaValLeuIleGluThrLeuLysValLeuGlyAl 87
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 TGATGTCAGATGGCTCTCCGCAACATCTCTCTACAAAGATACAGCGG 283
|:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87 agluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAla 104
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 CTGCTCTATCGTGTGCGGCCAACAGGACACAGCCAGAGAAAGCCAGCGGT 333
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
104 laAlaAlaIleAlaAla.....ThrGlyThr..... 112
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
334 ATCCCACTCTTCGCTCGGAAGGGGAAACACTCCAGAAATACCTGGGAGAA 383
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 ...ProValPheAlaValLysGlyGluThrLeuGluGluTyrTrpThrTy 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 CACATACCCGCTCTCACATGGCCAGATGGTCAAGGCCCCACAGCAGCGTTG 433
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
128 rhrAspGlnIlePheGlnTrpProAspGlyGluProSerAsnMetIleL 145
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 TCGATGATGGTGTGATGCTACACTCTCTCATCTCCAGGGGTTTCGAATTC 483
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 euAspAspGlyGlyAspAlaThrMetTyrIleLeuIleGlyAlaArgAla 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
484 GAAACAGCCGCTCTGTTCAGAGCCCAACAGAGCTGACACCTCGAATA 533
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GluAlaGlyGluAspValLeuSerAsnProGlnSerGlu.....GluG1 176
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
534 CGCTGCTGCTTCTGTACACTCAAGCAGAGTCTTCAACCAAGACAAAGAAC 583
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 uGluValLeuPheAlaGlnIleLysLysArgMetAlaAlaThrProGlyP 193
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
584 ACTGGCACACAGTTGCTGCGCGATGAACGGTGTTCGAGAGACAAACA 633
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
193 hePheThrLysGlnArgAlaAlaIleLysGlyValThrGluGluThrThr 209
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
634 ACAGTGTCCACCGCTCTACAGCTCGAAGAGGAGGCAAACTCCTCTT 683
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210 ThrGlyValAsnArgLeuTyrGlnLeuGlnLysLysGlyLeuLeuProPh 226
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
684 CCCAGCCATCAACGTCACAGCAGCTGTTCACAAAGTCCAAAGTTCGATAACA 733
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 eProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnL 243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
734 TCTACGGCTGCGCCACTCCCTTATCGATGGTATCAACCGCTCTCCGAT 783
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 ystyrglyCysLysGluSerLeuValAspGlyIleArgArgglyThrAsp 259
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
784 GTCATGATCGGGCAAGACAGCTCTCGTATGGTTACGGGTACGGCGATGTCG 833
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
260 ValMetMetAlaGlyLysValAlaValValCysglyTyrGlyAspValG1 276
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
834 CAAGGGCTCGCTCAATCCCTCGGTGCCCAAGGCGCTCGCTTATCATCA 883
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 yLysGlySerAlaGlnSerLeuAlaGlyAlaGlyAlaArgValLysValT 293
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
884 CAGAACTCGACCCCAATCTGGCTCTCCAGGCTGCCATGGAAGGCTACGAG 933
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 hrGluValAspProIleCysAlaLeuGlnAlaAlaMetAspGlyPheGlu 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
934 GTCCCGCGCATCGAGGAAGTCGTCAGAGTGTCTCGATATCTCTTCGTTACATG 983
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 ValValThrLeuAspAspAlaAlaSerThrAlaAspIleValValThrTh 326

```

984 CACAGGAAACTCGATATCATCTCTGTGACATGATGGCCAGATGAAGG 1033
 326 rThrGlyAsnLysAspValIleThrIleAspHisMetArgLysMetLysA 343
 1034 ATAAGGCTATTTCGTATACATCGGCACCTTCGATACGAAATGATACA 1083
 343 spMetCysIleValGlyAsnIleGlyHisPheAspAsnGluIleGlnVal 359
 1084 GATGGCTCATGAAATACCCAGCATCAACGACATCCCAATCAAGCCAGA 1133
 360 AlaAlaLeuArg.....AsnLeuLysTrpThrAsnValLysProGlu 373
 1134 APACGACATGTGGGAATCCAGATGCCACGCTATCTCTCTTCTGCTG 1183
 373 nValAspLeuIleGluPheProAspGlyLysArgLeuIleLeuLeuSerG 390
 1184 AGGCGCGCTTCTTACCTTGGCTGGCTACAGGTTCACCCATCTTCGTT 1233
 390 LuGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSerPheVal 406
 1234 ATGTCAATCTCATTCACAAACAGACACTCGCTCAGCTCGACCTCTACGA 1283
 407 MetSerAlaSerPheThrAsnGlnValLeuGlyGlnIleGluLeuPheTh 423
 1284 AAGAGAGGAAATCTCGAGAAGAGGTTTACACACTTCGGAAGCATCTCG 1333
 423 rArgThrAspAlaTyrLysAsnGluValTyrValLeuProLysHisLeuA 440
 1334 ATGCAAGACTCGCTCGCTCCCTCCACTCGGATCTCTCGATGTCCACTTACA 1383
 440 spLulysValAlaArgLeuHisLeuAspLysLeuGlyAlaLysLeuThr 456
 1384 AAGCTTACACAGAGGAGGCTGACTACATCAACGTTCCAGTTGAGGPTCC 1433
 457 ValLeuSerGluGluGlnAlaAlaTyrIleGlyValThrProGlnGlyPr 473
 1434 TTCAAGTCTGATGCTTACCGTTAT 1458
 473 oPheLysSerGluHisTyrArgTyr 481

seq_name: p1r1:A54040

seq_documentation_block:

adenosylhomocysteinase (EC 3.3.1.1) - malaria parasite (Plasmodium falciparum)

N;Alternate names: S-adenosyl-L-homocysteinase hydrolase

C;Species: Plasmodium falciparum

C;Date: 02-Aug-1994 #sequence_revision 04-Oct-1996 #text_change 09-Jun-2000

C;Accession: A54040

R;Creedon, K.A.; Rathod, P.K.; Wellem, T.E.

J. Biol. Chem. 269, 16364-16370, 1994

A;Title: Plasmodium falciparum S-adenosylhomocysteinase hydrolase. cDNA identification, p1

A;Reference number: A54040; MUID:94266832

A;Accession: A54040

A;Molecule type: mRNA

A;Residues: 1-479 <CRE>

A;Cross-references: GB:U07365; NID:g537431; PIDN:AAA21391.1; PID:g460639

C;Function:

C;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteinase to adeno

C;Superfamily: adenosylhomocysteinase

C;Keywords: NAD; thioether hydrolase

F;258-287/Region: beta-alpha-beta NAD nucleotide-binding fold

alignment_scores:

Quality: 1332.00 Length: 481

Ratio: 3.469 Gaps: 5

Percent Similarity: 79.834 Percent Identity: 53.430

alignment_block:

US-09-759-990-1 x A54040 ..

Align seg 1/1 to: A54040 from: 1 to: 479

40 AGAATTCGCCGACATCAACCTCCATGTTCTCGGCCGTAAAGAACCTTACCCCT 89
 7 LysValLysAspIleSerLeuAlaProPheGlyLysMetGlnMetGluI 23
 90 TGCTGAGAAGAAATCCAGTCTTATGTTCTTCGTGAGCGTTATTCGG 139
 23 eSerGluAsnGluMetProGlyLeuMetArgIleArgGluGluTyrGlyL 40
 140 CTCTTAAGCCATTGAAGGTGTCTAGAATCTCTGTGTTCCCTCCACATGACA 189
 40 ysAspGlnProLeuLysAsnAlaLysIleThrGlyCysLeuHisMetThr 56
 190 GTCCAGACACCGCTCTCTACGACACTCACAGCTCTTGTGCTCATGTT 239
 57 ValGluCysAlaLeuLeuIleGluThrLeuGlnLysLeuGlyAlaGlnI 73
 240 CAGATGGCTTCCTGCAACATCTCTACACAGATACACAGCGCTGCTG 289
 73 eArgTrpCysSerCysAsnIleTyrSerThrAlaAspTyrAlaAlaAla 90
 290 CTATCGTTGTGCGCCCAACAGCACACAGAGCCAGCCGCTATCCCA 339
 90 laVal.....SerThrLeuGlu.....AsnValThr 98
 340 GTCTTCGCCCTGGAAGGGGAAACACTCCCAAGAACTGGGAGAACACATA 389
 99 ValPheAlaTrpLysAsnGluThrLeuGluTyrTrpTrpCysValGlu 115
 390 CCGCGCTCTCACATGGCCAGATGGT.....CAAGGCCACACAGAGGTG 433
 115 uSerAlaLeuThrTrpGlyAspGlyAspAsnGlyProAspMetIleV 132
 434 TCGATGATGTTGTGATGCTACCTCTCATCTCTCAAGGGCTTCGAA... 480
 132 alaAspAspGlyGlyAspAlaThrLeuLeuValHisLysGlyValGluTyr 148
 481TTCGAAACAGCGGTCTCTCCAGAGCCCAACAGAGCTCA 521
 149 GluLysLeuTyrGluGluLysAsnIleLeuProAspProGluLysAlaL 165
 522 CAACCTCGAATACCGCTGCTTCTGTCTACACTCAAGCAGGTCTTCAACC 571
 165 sAsnGluGluArgCysPheLeuThrLeuLeuLysAsnSerIleLeuL 182
 572 AAGACAGAACCTACGACACAGATGCTGCGCGCATGACCGGTGTTTCC 621
 182 ysAsnProLysLysTrpThrAsnIleAlaLysLysIleIleGlyValSer 198
 622 GAAGAGACAAACAGGTGTCCACGCTCTACCAGCTCGAGAGGAGGG 671
 199 GluGluThrThrThrGlyValLeuArgLeuLysLysMetAspLysGlnAs 215
 672 CAAACTCTCTTCCAGCCATCAACGCTCAACGACGCTGTACAAAGTCCA 721
 215 nGluLeuLeuPheThrAlaIleAsnValAsnAspAlaValThrLysGlnL 232
 722 AGTTCTGATACATCTACGGTCCGCCACTCCCTTATCGATGATCAAC 771
 232 ystAspAsnValTyrGlyCysArgHisSerTyrThrAspGlyLeuMet 248
 772 CGTCTTCCGATGTCATCGCGGCAAGACAGCTCTCGTCATGGGTGA 821
 249 ArgAlaThrAspPheLeuIleSerGlyLysIleValIleCysGlyTy 265
 822 CGCGATGTCCGCAAGGGCTCGCTCAATCCCTCGGTGCCAAGGCGCTC 871
 265 rGlyAspValGlyLysGlyCysAlaSerSerMetLysGlyLeuGlyAla 282
 872 GCGTTATCATCAGAACTCCACCCAAATCTGGCTCTCCAGGCTGCCATG 921
 282 rgValTyrIleThrGluIleAspProIleCysAlaIleGlnAlaValMet 298
 922 GAAGGCTACCAAGTCCGCGCATCGAGGAAGTCTCAAGGATGTCGATAT 971

```

|||||.....||| .....|||.....||| .....|||
299 GluGlyPheAsnValValThrLeuAspGluIleValAspLysGlyAspPh 315
972 CTTCCTTACATGCACGAAGAACTGCATATCATCTCTGTTCACATGATGG 1021
315 ePheIleThrCysThrGlyAsnValAspValIleLysLeuGluHisLeuL 332
1022 CCCAGATGAAGGATTAAGCTATTGTCGTAACATCGGCCACTTCGATAAC 1071
332 euLysMetLysAsnAsnAlaValValGlyAsnIleGlyHisPheAspAsp 348
1072 GAAATGTATACAGATGGCTCATGAATACCCAGCATCAAGCACATCCC 1121
349 GluIleGlnValAsnGluLeuPheAsnTyrLysGlyIleHisIleGluAs 365
1122 AATCAACCCAGAAATPACGACATGTGGAAATCCAGATGCCACCGCTATCC 1171
365 nValLysProGlnValAspArgIleThrLeuProAsnGlyAsnLysIleI 382
1172 TCCTCTCTGCTGAGGCGCGCTCTTACCTTGGCTGCCTACAGTCCAC 1221
382 leValLeuAlaArgGlyArgLeuLeuAsnLeuGlyCysAlaThrGlyHis 398
1222 CATCTCTTCTTATGTCATATGTCATTCACAAACACAGACACTCGCTCAGCT 1271
399 ProAlaPheValMetSerPheSerPheCysAsnGlnThrPheAlaGlnLe 415
1272 CGACCTCTACGAAAGAGA.....GGAATCTCGAGAAGAGTTTACA 1315
415 uAspLeuTrpGlnAsnLysAspThrAsnLysTyrGluAsnLysValTyrL 432
1316 CACTTCCGAGCATCTCGATGAAGAACTGCTCGCTCCCTCCACCTCGGATCT 1365
432 euLeuProLysHisLeuAspGluLysValAlaLeuTyrHisLeuLysLys 448
1366 CTCGATGTCCTCATCAAGCTTACACAGAGCAGGCTGACTACATCAA 1415
449 LeuAsnAlaSerLeuThrGluLeuAspAspAsnGlnCysGlnPheLeuG 1465
1416 CTTCCAGTTGAGGTCCTTACAAAGTCTGATGCTTACCGTTAT 1458
465 yValAsnLysSerGlyProPheLysSerAsnGluTyrArgTyr 479
seq_name: piri:A26583

```

```

seq_documentation_block:
adenosylhomocysteine (EC 3.3.1.1) - rat
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Nov-1988 #sequence_revision 04-Oct-1996 #text_change 18-Jun-1999
R;Ogawa, H.; Gomi, T.; Mueckler, M.M.; Fujioka, M.; Backlund Jr., P.S.; Aksamit, R.R.; U
Proc. Natl. Acad. Sci. U.S.A. 84, 719-723, 1987
A;Title: Amino acid sequence of S-adenosyl-L-homocysteine hydrolase from rat liver as de
A;Reference number: A26583; MUID:87118240
A;Accession: A26583
A;Molecule type: mRNA
A;Residues: 1-432 <OGA>
A;Cross-references: GB:M15185; NID:g202803; PIDN:AAAA0705.1; PID:g202804
R;Gomi, T.; Ogawa, H.; Fujioka, M.
J. Biol. Chem. 261, 13422-13425, 1986
A;Title: S-Adenosylhomocysteine from rat liver. Amino acid sequences of the peptides c
A;Reference number: A92560; MUID:87008564
A;Accession: A26191
A;Molecule type: protein
A;Residues: 76-94 <GOM>
A;Accession: B26191
A;Molecule type: protein
A;Residues: 104-121 <GO2>
A;Note: in a manner saturably inhibited by adenosine (a product) or adenine (a competitor
-79 and Cys-113, suggesting a role for these residues at the active site; however, neith
R;Merta, A.; Aksamit, R.R.; Kasir, J.; Cantoni, G.L.
Eur. J. Biochem. 229, 575-582, 1995

```

```

A;Title: The gene and pseudogenes of rat S-adenosyl-L-homocysteine hydrolase.
A;Reference number: S69333; MUID:95262723
A;Accession: S69333
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <MEP>
A;Cross-references: EMBL:U14937; NID:g1223843; PIDN:AAA92043.1; PID:g1185363
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
C;Genetics:
A;Introns: 10/1; 73/3; 99/1; 149/1; 186/3; 256/1; 285/2; 324/3; 389/3
C;Function:
A;Description: catalyzes the reversible hydrolysis of S-adenosyl-L-homocysteine to ad
C;Superfamily: adenosylhomocysteinease
C;Keywords: homotetramer; NAD; thioether hydrolase
F;214-243/Region: beta-alpha-beta NAD nucleotide-binding fold
F;79,113/Active site: Cys #status predicted

```

alignment_scores:

```

Quality: 1289.00 Length: 490
Ratio: 3.662 Gaps: 6
Percent Similarity: 71.837 Percent Identity: 56.122

```

alignment_block:

```

US-09-759-990-1 x A26583

```

```

Align seg 1/1 to: A26583 from: 1 to: 432

```

```

1 ATGGCTTCAATCACTACTGTGTCTCCATTCGAGTACAGAAATTCGCGA 50
||||| ||| ||| |||.....|||
1 MetAlaAspLysLeuPro.....TyrLysValAlaAs 11
51 CATCAACTCCATGTTCTCGCCGTAAGAACTTACCTTGTGAGAAAG 100
|||||.....||| ||| |||.....|||
11 PileGlyLeuAlaAlaTrpGlyArgLysAlaLeuAspIleAlaGluAsn 28
101 AATGCCAGGTCTTATGTTCTTCTGAGCGTTATTCGCTTCTTAAGCCA 150
|||||.....||| ||| |||.....|||
28 LuMetProGlyLeuMetArgMetArgGluMetTyrSerAlaSerLysPro 44
151 TTGAAGGTCACAACTCTGTTCCCTCCACATGACAGTCCAGACAGC 200
|||||.....||| ||| |||.....|||
45 LeuLysGlyAlaArgIleAlaGlyCysLeuHisMetThrValGluThrAl 61
201 CGTCCTCATCGACACTCACAGCTCTTGGTGTGATGTCAGATGGCTT 250
|||||.....||| ||| |||.....|||
61 aValLeuIleGluThrLeuValAlaLeuGlyAlaGluValArgTrpSerS 78
251 CTGGCAACATCTTCTTACACAGATACAGCCGCTGCTGCTATCCTTGT 300
|||||.....||| ||| |||.....|||
78 exCysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAla... 93
301 GGCCCAACAGGCACACAGAGAGCCAGCCGCTATCCAGTCTTCGCTG 350
|||||.....||| ||| |||.....|||
94 .....LysAlaGlyIleProValPheAlaTr 102
351 GAAGGCGCAACACTCCAGAAATACGGAGAACACATACCGCGCTCTCA 400
|||||.....||| ||| |||.....|||
102 PlysGlyGluThrAspGluGluTyrLeuTyrCysIleGluGlnThrLeuH 119
401 CATGCCAGATGTCAGGCCACAGCAGGTTGTCGATGATGGTGTAT 450
|||||.....||| ||| |||.....|||
119 IsPheLysAspGlyPro...LeuAsnMetIleLeuAspGlyGlyAsp 134
451 GCTACACTCTCATCTCCCAAGGCTTCGAATTCGAACACGCCGCTGT 500
|||||.....||| ||| |||.....|||
135 LeuThrAsnLeuIle..... 139
501 TCCAGAGCCACAGAGCTGACAACTCGAATACCGTGGTCTTCTTCTA 550
139 ..... 139
551 CACTCAAGCAGGTCTTCAACCAAGACAAGAACCCACTGGCACACA..... 594

```

426 ysProAspHisTyrArgTyr 432

seq_name: pir2:T40763

seq_documentation_block:
adenosylhomocysteinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T40763
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21948
A;Accession: T40763
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-433 <LYN>
A;Cross-references: EMBL:AL02072; PIDN:CAAL7833.1; GSPDB:GN00067; SPDB:SPBC8D2.18C
A;Experimental source: strain 972h-; cosmid c8D2
C;Genetics:
A;Gene: SPDB:SPBC8D2.18C
A;Map position: 2
C;Superfamily: adenosylhomocysteinase

alignment_scores:
Quality: 1274.00 Length: 475
Ratio: 3.559 Gaps: 5
Percent Similarity: 75.368 Percent Identity: 54.526

alignment_block:

US-09-759-990-1 x T40763 ..

Align seg 1/1 to: T40763 from: 1 to: 433

```
37 TACAGAAATCCGACATCAACCTCCATGTTCTCGCGCTAGGAACATTAC 86
|||||
6 TyrLysValAlaAspLysSerLeuAlaAlaPheGlyArgLysGluLeuGl 22
|||||
87 CCTGCTGAGAAAGAAATGCCAGGCTTATGTTCTTCTGCTGAGCGCTATT 136
|||||
22 uileAlaGluasnGluMetProGlyLeuileAlaValArgGluLysTyrA 39
|||||
137 CCGCTTCTTAAGCCATTGAAGGTTGTCAGAAATCTCTGTTCTCCCTCCACATG 186
|||||
39 laLysSerGlnProLeuLysGlyAlaArgIleAlaGlyLysLeuHisMet 55
|||||
187 ACAGTCAGACAGCGCTCTCATCGAGACACTCACAGCTCTGTGCTGA 236
|||||
56 ThrIleGlnThrAlaValLeuileGluThrLeuValAlaLeuGlyAlaGl 72
|||||
237 TGTGAGATGGGCTTCTCTCAACATCTTCTTACACAAGATACAGCGGCTG 286
|||||
72 uValThrTrpSerSerCysAsnIleTyrSerThrGlnAspHisAlaAla 89
|||||
287 CTGCTATCTGTTGTCGCGCCACAGGACAGGACAGAGCCAGCGGTATC 336
|||||
89 laAlaIleAlaAla.....ThrGlyVal 96
|||||
337 CCAGTCTTCGCTGGAAGGGCGAAACACTCCAGAAATAC...TGGGAGAA 383
|||||
97 ProValPheAlaTyrPlyGlyGluThrGluGluLutyrLeuTrpCysIl 113
|||||
384 CACATACCGCGCTCTCATATGGCAGATGTTCAAGGCCACACAGCGGTTG 433
|||||
113 egluGlnGlnLeuLysSerPheProSerGlyLysProLeuAsnMetIleL 130
|||||
434 TCGATCATGCTGCTGATGCTACACTCTCTCATCTCCAAAGGCTTCGAATC 483
|||||
130 euAspPlyGlyLysPheLeuThrAlaLeuValHisGluArgHis..... 144
|||||
484 GAAACAGCGGCTGCTCTTCAGAGCCACAGAGAGAGCTGACAACTCGAATA 533
|||||
145 .....ProGlu..... 146
```

534 CCCTCGCTTCTTCTGCTACACTCAAGCAGGTCTTCAACCAAGACAAGAAC 583
146 146
584 ACTGGCACACAGTTCGTGGCGGATGACGGTGTTCGGAAGACAACA 633
147 147
634 ACAGGTGTCCACGCCCTCTACCACTCGAGAGGAGGAGGAACTCCTCTT 683
160 ThrGlyValHisAsnLeuTyrLysMetPheLysGluAsnLysLeuLysVa 176
684 CCAGCCATCAAGCTCAACGACCTGTTCACAAAGTCCAAGTTCGATAACA 733
176 IProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspAsnL 193
734 TCTACGGCTGCGGCCCTCCCTTATCGATGATCAACCGTTCCTCCGAT 783
193 euPheGlyCysLysGluSerLeuValAspGlyIleLysArgAlaThrAsp 209
784 GTCATGATCGCGGCGGCAAGACAGCTCTCGTCATGGGTACGGCGATGTCGG 833
210 ValMetIleAlaGlyLysValAlaValAlaGlyPheGlyAspValGI 226
834 CAAGGCTGCGCTCAATCCCTCGGTGGCCCAAGCGCTCGGTTCATCATCA 883
226 YLysGlyCysSerThrSerLeuArgSerGlnGlyAlaArgValIleValT 243
884 CAGAACTCGACCAATCTCTCTGTCACATGATGCCCGCCACATGAAG 1033
276 rThrGlyCysArgAspIleIleArgGlyGluHisPheAsnGluMetLysG 293
1034 ATAAGGCTATTGTCGGTAACTTCGACATCGGCCACTTCGATAACGAAATGATACA 1083
293 luAspSerIleValCysAsnIleGlyHisPheAspValGluIleAspVal 309
1084 GATGGCTCATGAAATACCCAGGCATCAAGCACATCCCAATCAAGCCAGA 1133
310 ...AlaTrpLeuLysAlaAsnAlaLysAspValValAsnIleLysProGI 325
1134 ATAGCACATGTGGAAATCCAGATGGCCAGCTATCCCTCTTCTGCTG 1183
325 nValAspArgTyrGluLeuLysAsnGlyArgHisIleIleLeuLeuAla 342
1184 AGGCGCGCTTCTTAACCTTGGCTGCGGTACAGGTCAAGTTCCTTTCGTT 1233
342 spGlyArgLeuValAsnLeuGlyCysAlaThrGlyHisProSerPheVal 358
1234 ATGTCATGTCATCAACCAAGACACTCGCTCAGCTCGACCTCTACGA 1283
359 MetSerCysSerPheThrAsnGlnValLeuAlaGlnIleAlaLeuTrpTh 375
1284 AAAGAGAGGAATCTCGAGAGAAGGTTTACACACTTCCGAGCATCTCG 1333
375 rAspAsnThrSerTyrProLeuGlyValHisMetLeuProLysLysLeuA 392
1334 ATGAAGAAGTCGCTCCCTCCAGCTCTCTCGATGTCGATGTCACCTTACA 1383
392 spGluGluValAlaArgLeuHisLeuGlyLysLeuGlyValLysLeuThr 408
1384 AGCTTACAGCAAGCGGTGACTACATCAACGTTCCAGTTGAGGTC 1433
409 ThrLeuThrSerValGlnSerAspTyrLeuGlyIleProValAspGlyPr 425

1434 TTACAAGTCTGATGCTTACCGTTAT 1458
|||||:|||||
425 OTYrLysAlaAspHisTyrArgTyr 433

us-09-759-990-1.n2p.rpr

Mon Aug 19 15:16:21 2002

OM of: US-09-759-990-1 to: A_Geneseq_032802:* out_format : pfs
 Date: Aug 17, 2002 1:46 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frameat_n2p.model -DEV=rlp
 -Q/cgn2_1/USPTO_spool_US0975990/runat_13082002_153924_28014/app_query.fasta_1.1541
 -DB=A_Geneseq_032802 -QFMT=FASTA -SUFFIX=n2p.rag -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=BIOSUM62
 -TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
 -USER=US0975990_0CGN1_1.122 -NGPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLUPY -WAIT -THREADS=1

Search information block:

Query: US-09-759-990-1
 Query length: 1461
 Database: A_Geneseq_032802:*
 Database sequences: 747574
 Database length: 111073796
 Search time (sec): 102.110000

score_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.1e-128	49			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.8e-127	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.8e-127	50			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.7e-126	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	3.2e-126	51			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	3.2e-126	46			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	6.8e-125	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	3.0e-125	49			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	7.7e-124	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.1e-123	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.6e-123	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	8.1e-122	45			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.2e-120	46			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.2e-120	48			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.0e-120	47			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.0e-120	47			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	8.4e-119	44			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.5e-108	43			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.5e-108	43			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.5e-108	43			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	6.4e-103	43			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	6.4e-103	43			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.6e-87	50			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	7.0e-85	52			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	9.3e-83	61			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.3e-82	61			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.0e-76	46			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.1e-56	42			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.4e-52	45			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.8e-39	23			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.1e-29	24			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	1.9e-16	81			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	2.81	97			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	4.7e-16	95			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	266.50	31			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	376.70	13			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	314.25	41			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	208.00	23			
/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873	302.62	56			

/SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAU34408 + 205.00 300.72 2.7e-09
 /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAU34408 + 197.00 277.36 2.2e-08
 /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:ABG28410 + 193.50 254.25 9.4e-08
 /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAU95559 + 177.50 236.40 1.8e-06
 /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAU95556 + 177.50 235.20 1.9e-06

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA87873

seq_documentation_block:

ID_AA87873 standard; Protein; 495 AA.
 AC_AA87873;
 DT 06-OCT-2000 (first entry)
 DE M. tuberculosis antigen TB54 protein.
 KW Tuberculosis; TB; antigen; vaccine; diagnosis; somatic; tuberculostatic;
 KW infection; interferon-gamma; IFN-gamma; protective immunity; therapy;
 KW delayed type hypersensitivity response; TB54.
 OS Mycobacterium tuberculosis.
 PN WO200021983-A2.
 PD 20-APR-2000.
 PF 08-OCT-1999; 99WO-DK00538.
 PR 08-OCT-1998; 98DK-0001281.
 PR 21-JAN-1999; 99US-0116673.
 XX (STAT-) STATENS SERUM INST.
 XX Andersen P, Weldingh K, Hansen CV, Florio W, Okkels LM;
 XX Skjot RLV, Rosenkrands I;
 XX WPI; 2000-317931/27.
 XX N-PSDB; AAA39567.
 XX Novel polypeptide of somatic protein extract useful as vaccine against
 XX virulent Mycobacterium infection, isolated from cell wall, cell
 XX membrane and cytosol .
 XX Claim 1; Page 96-97; 126pp; English.
 XX This invention describes a novel polypeptide (pp) of somatic proteins
 XX extract (I) which have tuberculostatic activity. (I) or their subsequence
 XX has at least one of the following properties: (a) the pp induces an in
 XX vitro recall response, or an in vitro response, during primary infection
 XX with virulent Mycobacterium, determined by a release of interferon
 XX (IFN)-gamma, (b) pp induces a protective immunity, determined by
 XX vaccinating an animal with pp and an adjuvant, three times at two weeks
 XX intervals, (c) pp induces an in vitro response, or in vitro recall
 XX response, determined by release of IFN-gamma of at least 1000 pg/ml or
 XX 500 pg/ml, respectively, from Peripheral Blood Mononuclear Cells (PBMC)
 XX withdrawn from TB patients, or PPD positive individuals, 6 months after
 XX diagnosis, (d) pp induces a specific antibody response in a TB patient,
 XX as determined by enzyme linked immunosorbent assay (ELISA) technique or
 XX a western blot, (e) pp induces a positive delayed type hypersensitivity
 XX (DTH) response, determined by intradermal injection. (I) and (II) are
 XX useful in preparing a prophylactic or therapeutic vaccine as a vaccine
 XX for induction of a protective or generation of an immune response in a
 XX mammal against infection with a virulent Mycobacterium. (I) and (II) are
 XX also useful as diagnostic reagent for the diagnosis of a virulent
 XX Mycobacterium infection. The vaccine of the invention induces efficient
 XX immunological memory, providing long term protection against TB. This
 XX sequence represents a Microbacterium tuberculosis TB54 antigen
 XX described in the invention.
 XX Sequence 495 AA;

alignment_scores:

Quality: 1550.50 Length: 477
 Ratio: 3.945 Gaps: 3
 Percent Similarity: 82.390 Percent Identity: 62.264

alignment_block:
 US-09-759-990-1 x AY87873 ..

Align seg 1/1 to: AY87873 from: 1 to: 495

```

34 GAGTACAGAAATGCCGACATCAACCTCATCTCTCGCCGCGTAAGGAAC 83
20 AppPhyLysIleAlaAspLeuSerLeuAlaAspPheGlyArgLysGluLe 36
84 TACCCCTGTGTGAGAGAAATGCCAGGTCTTATGTTCGTTCGTGAGCGTT 133
36 uArgIleAlaGluHisGluMetProGlyLeuMetSerLeuArgGluT 53
134 ATTCCGCTTCTAAGCCATTGAAGGGTGTCAAGATCTCTGTTCCCTCCAC 183
53 yAlaGluValGlnProLeuLysGlyAlaArgIleSerGlySerLeuHis 69
184 ATGACAGTCCAGACAGCCGCTCATCTCGACACTCACAGCTCTTGTGTC 233
70 MetThrValGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 86
234 TGATGTCAGATGGCTTCCTGCAACATCTTCTTACACAGATACAGCCG 283
86 aGluValArgTrpAlaSerCysAsnIlePheSerThrGlnAspHisAla 103
284 CTGCTGCTATCTGTCGCGCCACACAGGCACACAGAGCCAGCCGCT 333
103 laAlaAlaValValGlyProHisGlyThrProAspGluProLysGly 119
334 ATCCAGTCTTCCCTCGGAAGGCGCAACACTCCAGAAATCTGGGAAA 383
120 ValProValPheAlaTrpLysGlyGluThrLeuGluGluTrpAl 136
384 CACATACCCGCTCTCAGATGCCAGATGTCAAGGCCA...CAGCAGG 430
136 aAlaGluGlnMetLeuThrTrpProAspProAspLysProAlaAsnMetI 153
431 TTGTCGATGATGTGTGTGATGTCTACACTCTCTCATCTCCAGGGCTCGAA 480
153 leuAspAspGlyGlyAspAlaThrMetLeuValLeuArgGlyMetGln 169
481 TTCGAACACCGCGTCTGTTCAGAGCCCAACAGAGCTGACAACTCGA 530
170 TyrGluLysAlaGlyValProAlaGluGluAspAspProAlaG 186
531 ATACCGCTCGGTTCTGTGTACACTCAAGCAGGTCTTCAACCAACAGAGA 580
186 utrLysValPheLeuAsnLeuLeuArgThrArgPheGluThrAspLysA 203
581 ACCACTGGCACAGATGTGTCGCGGATGAACGTGTTCCGAGAGAGACA 630
203 splYstrPthrLysIleAlaGluSerValLysGlyValThrGluGluThr 219
631 ACAACAGGTGTCCACCGCTCTACAGCTCGAGAGGAGGCGCAACTCCT 680
220 ThrThrGlyValLeuArgLeuTyrGlnPheAlaAlaGlyAspLeuAl 236
681 CTTCGAGCATCAACGTCAACGACGCTGTATACAAAGTCCAGTTCGATA 730
236 apHeProAlaIleAsnValAsnAspSerValThrLysSerLysPheAspA 253
731 ACATCTACGCTCGCCGCTCTCTTATCGATGTATCAACCGCTCTCC 780
253 snLysTyrGlyThrArgHisSerLeuIleAspGlyIleAsnArgGlyThr 269
781 GATGTCATGATCGCGGCACAGACAGCTCTCTGTCATGGGTTACGCGCATGT 830
270 AspAlaLeuIleGlyGlyLysValLeuIleCysGlyTyrGlyAspVa 286

```

```

831 CGCAAGGGTGCCTCAATCCCTCCGTGGCCAGCGCTCGCGTTATCA 880
286 lGlyLysGlyCysAlaGluAlaMetLysGlyGlnGlyAlaArgValSerV 303
881 TCACAGAACTCGACCCCAATCTGCGCTCTCCAGGCTGCCATGGAAGCTAC 930
303 alThrGluIleAspProIleAsnAlaLeuGlnAlaMetMetGluGlyPhe 319
931 CAGTCCGCGCATCGAGGAAGTGTCAAGGATGTCTGATATCTTCTGTTAC 980
320 AspValValThrValGluGluAlaIleGlyAspAlaAspIleValValTh 336
981 ATGACAGAACTCGCATATCATCTCTGTTGATGATGATGATGATGATGATG 1030
336 rAlaThrGlyAsnLysAspIleIleMetLeuGluHisIleLysAlaMetL 353
1031 AGGATAGGCTATTGTGCGTAACATCGGCCTCTGATGATGATGATGATGAT 1080
353 yAspHisAlaIleLeuGlyAsnIleGlyHisPheAspAsnGluIleAsp 369
1081 ACAGATGGCTCATGAATACCCAGGCATCAAGCACATCCCCAATCAAGCC 1130
370 MetAlaGlyLeuGluArg...SerGlyAlaThrArgValAsnValLysPr 385
1131 AGAATACGACATGTGGAAATCCAGAT...GGCCACGCTATCTCTCTTC 1177
385 oGlnValAspLeuTrpPheGlyAspThrGlyArgSerIleIleValL 402
1178 TTGCTGAGGCGCGCTTCTTAACCTTGCTGCTGCTGCTGCTGCTGCTGCT 1227
402 eusGlyGlyArgLeuLeuAsnLeuGlyAsnAlaThrGlyHisProSer 418
1228 TTGCTTATGTCAATGTCTTCAAAACACAGACACTCGCTCAGCTCGACCT 1277
419 PheValMetSerAsnSerPheAlaAsnGlnThrIleAlaGlnIleGluLe 435
1278 CTAGAAAGAGAGAAATCTCGAGAGAGAGGTTTACACACTTCCGAGC 1327
435 utrPThrLysAsnAspGluTyrAspAsnGluValTyrArgLeuProLysH 452
1328 ATCTCGATGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1377
452 isLeuAspGluLysValAlaArgIleHisValGluAlaLeuGlyHis 468
1378 CTACAAAGCTTACACAGAGAGCTGACTACATCAACAGTTTCCAGTTGA 1427
469 LeuThrLysLeuThrLysGluGlnAlaGluTyrLeuGlyValAspValG 485
1428 GGTCTTACAGTCTGATGTTACCGTTAT 1458
485 uGlyProTyrLysProAspHisTyrArgTyr 495

```

seq_name: /SDSS/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG40086

seq_documentation_block:

ID AAG40086 standard; Protein; 485 AA.

XX AAG40086;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 49691.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
 PF 99US-0121825.
 XX 25-FEB-1999;
 PR 99US-0123180.
 PR 05-MAR-1999;
 PR 09-MAR-1999;
 PR 23-MAR-1999;
 PR 25-MAR-1999;
 PR 29-MAR-1999;
 PR 01-APR-1999;
 PR 08-APR-1999;
 PR 16-APR-1999;
 PR 19-APR-1999;
 PR 21-APR-1999;
 PR 23-APR-1999;
 PR 28-APR-1999;
 PR 30-APR-1999;
 PR 30-APR-1999;
 PR 04-MAY-1999;
 PR 05-MAY-1999;
 PR 06-MAY-1999;
 PR 07-MAY-1999;
 PR 11-MAY-1999;
 PR 14-MAY-1999;
 PR 14-MAY-1999;
 PR 14-MAY-1999;
 PR 18-MAY-1999;
 PR 19-MAY-1999;
 PR 20-MAY-1999;
 PR 21-MAY-1999;
 PR 24-MAY-1999;
 PR 25-MAY-1999;
 PR 27-MAY-1999;
 PR 28-MAY-1999;
 PR 01-JUN-1999;
 PR 03-JUN-1999;
 PR 04-JUN-1999;
 PR 07-JUN-1999;
 PR 08-JUN-1999;
 PR 10-JUN-1999;
 PR 10-JUN-1999;
 PR 16-JUN-1999;
 PR 17-JUN-1999;
 PR 17-JUN-1999;
 PR 18-JUN-1999;
 PR 18-JUN-1999;
 PR 18-JUN-1999;
 PR 18-JUN-1999;
 PR 18-JUN-1999;
 PR 21-JUN-1999;
 PR 21-JUN-1999;
 PR 22-JUN-1999;
 PR 23-JUN-1999;
 PR 23-JUN-1999;
 PR 24-JUN-1999;
 PR 29-JUN-1999;
 PR 30-JUN-1999;
 PR 01-JUL-1999;
 PR 01-JUL-1999;
 PR 02-JUL-1999;
 PR 06-JUL-1999;
 25-FEB-2000; 2000EP-0301439.
 99US-0121825.
 99US-0123180.
 99US-0123548.
 99US-0125788.
 99US-0126264.
 99US-0126785.
 99US-0127462.
 99US-0128234.
 99US-0128714.
 99US-0129845.
 99US-0130077.
 99US-0130449.
 99US-0130510.
 99US-0130891.
 99US-0131449.
 99US-0132048.
 99US-0132407.
 99US-0132484.
 99US-0132485.
 99US-0132486.
 99US-0132487.
 99US-0132863.
 99US-0134256.
 99US-0134218.
 99US-0134219.
 99US-0134221.
 99US-0134370.
 99US-0134768.
 99US-0134941.
 99US-0135124.
 99US-0135353.
 99US-0135629.
 99US-0136021.
 99US-0136392.
 99US-0136782.
 99US-0137222.
 99US-0137528.
 99US-0137502.
 99US-0137724.
 99US-0138094.
 99US-0138540.
 99US-0138847.
 99US-0139119.
 99US-0139452.
 99US-0139453.
 99US-0139452.
 99US-0139453.
 99US-0139454.
 99US-0139455.
 99US-0139456.
 99US-0139457.
 99US-0139458.
 99US-0139459.
 99US-0139460.
 99US-0139461.
 99US-0139462.
 99US-0139463.
 99US-0139750.
 99US-0139763.
 99US-0139817.
 99US-0139859.
 99US-0140353.
 99US-0140354.
 99US-0140695.
 99US-0140823.
 99US-0140991.
 99US-0141287.
 99US-0141842.
 99US-0142154.
 99US-0142055.
 99US-0142390.
 08-JUL-1999;
 09-JUL-1999;
 12-JUL-1999;
 13-JUL-1999;
 14-JUL-1999;
 15-JUL-1999;
 16-JUL-1999;
 16-JUL-1999;
 19-JUL-1999;
 19-JUL-1999;
 19-JUL-1999;
 19-JUL-1999;
 20-JUL-1999;
 20-JUL-1999;
 20-JUL-1999;
 21-JUL-1999;
 21-JUL-1999;
 21-JUL-1999;
 22-JUL-1999;
 22-JUL-1999;
 22-JUL-1999;
 23-JUL-1999;
 23-JUL-1999;
 26-JUL-1999;
 27-JUL-1999;
 27-JUL-1999;
 28-JUL-1999;
 28-JUL-1999;
 02-AUG-1999;
 02-AUG-1999;
 03-AUG-1999;
 04-AUG-1999;
 05-AUG-1999;
 05-AUG-1999;
 06-AUG-1999;
 06-AUG-1999;
 09-AUG-1999;
 09-AUG-1999;
 11-AUG-1999;
 12-AUG-1999;
 13-AUG-1999;
 13-AUG-1999;
 16-AUG-1999;
 17-AUG-1999;
 18-AUG-1999;
 20-AUG-1999;
 20-AUG-1999;
 23-AUG-1999;
 23-AUG-1999;
 25-AUG-1999;
 26-AUG-1999;
 27-AUG-1999;
 27-AUG-1999;
 30-AUG-1999;
 31-AUG-1999;
 01-SEP-1999;
 07-SEP-1999;
 10-SEP-1999;
 13-SEP-1999;
 15-SEP-1999;
 16-SEP-1999;
 20-SEP-1999;
 22-SEP-1999;
 23-SEP-1999;
 99US-0142803.
 99US-0142920.
 99US-0142977.
 99US-0143542.
 99US-0143624.
 99US-0144005.
 99US-0144085.
 99US-0144086.
 99US-0144325.
 99US-0144331.
 99US-0144332.
 99US-0144333.
 99US-0144334.
 99US-0144335.
 99US-0144352.
 99US-0144632.
 99US-0144884.
 99US-0144814.
 99US-0145086.
 99US-0145088.
 99US-0145085.
 99US-0145087.
 99US-0145089.
 99US-0145192.
 99US-0145145.
 99US-0145218.
 99US-0145224.
 99US-0145276.
 99US-0145913.
 99US-0145918.
 99US-0145919.
 99US-0145951.
 99US-0146386.
 99US-0146388.
 99US-0146389.
 99US-0147038.
 99US-0147204.
 99US-0147302.
 99US-0147192.
 99US-0147260.
 99US-0147303.
 99US-0147416.
 99US-0147493.
 99US-0147935.
 99US-0148171.
 99US-0148319.
 99US-0148341.
 99US-0148565.
 99US-0148684.
 99US-0149368.
 99US-0149175.
 99US-0149426.
 99US-0149722.
 99US-0149723.
 99US-0149929.
 99US-0149902.
 99US-0149930.
 99US-0150566.
 99US-0150884.
 99US-0151065.
 99US-0151066.
 99US-0151080.
 99US-0151303.
 99US-0151438.
 99US-0151930.
 99US-0152363.
 99US-0152363.
 99US-0153070.
 99US-0153758.
 99US-0154018.
 99US-0154039.
 99US-0154779.
 99US-0155139.
 99US-0155486.

```

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 1537.00      Length: 487
Ratio: 3.901          Gaps: 4
Percent Similarity: 80.903      Percent Identity: 62.012

alignment_block:
US-09-759-990-1 x AAG40086 ..

Align seg 1/1 to: AAG40086 from: 1 to: 485

19 ACTGGTGCTCCATTCGAGTACAGNAATTCGCGACATCAACCTCCATGTTCT 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerglyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

69 CGCGCGTAAGGAAGCTTACCTGCTGAGAGGAATCCAGGCTCTATGG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluLeuAlaGluValGluMetProGlyLeuMetA 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

119 TTTCTGCTGAGCGTTATTCCTCTTAAAGCATTCGAGGGTGTGAAATC 168
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
41 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 57
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

169 TCTGTTCCCTCCACATGACAGTCCAGACGCGCTCTCATCGAGACACT 218
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuLeuGluThrLe 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

219 CACAGCTCTTGGTGCTGATGTCAGATGGGCTCTTCGCAACATCTCTTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
74 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerT 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

269 CACAAGATACAGCCGCTCTGCTATCTGTTGTCGGCCCAACAGCGCACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 hrGlnAspHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

319 GAGAGCCAGCGGTATCCAGTCTTCGCTGGAAGGGCGGAACACTCC 368
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

... 10 1514 1555, 3303 0134/08.

PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	11-AUG-1999;	99US-0148341.
PR	12-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	13-AUG-1999;	99US-0149368.
PR	16-AUG-1999;	99US-0149175.
PR	17-AUG-1999;	99US-0149426.
PR	18-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
alignment_scores:		
Quality: 1537.00		Length: 487
Ratio: 3.901		Gaps: 4
Percent Similarity: 80.903		Percent Identity: 62.012
alignment_block:		
US-09-759-990-1 x AAG40085 ..		
Align seg 1/1 to: AAG40085 from: 1 to: 508		
19	ACTGGTCTCATTCGAGTACAGAAATGCCGACATCAACCTCCATGTTCT	68
31	ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh	47
69	CGCCGTAAGGAACCTTACCTTGCTGAGAAGAAATGCCAGGTCTTATGG	118
47	eglyArgLeuGluLeuGluAlaGluValGluMetProGlyLeuMeta	64
119	TTCCTGCTGACGGTTATTCGGCTTCTAAGCCATTGAAGGGTGTCAATC	168
64	laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile	80
169	TCTGTTCCCTCCACATGACAGTCCAGACGCGCTCTCATCGAGACACT	218
81	ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe	97
219	CACAGCTCTTGGTGTGATGTCAGATGGCTTCTCTCAACATCTTCTTA	268
97	uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSert	114
269	CACAAGATACAGCCGCTGCTGTATGCTGTCGGCCCAACAGGCACACCA	318
114	hrGlnAspHisAlaAlaAlaIleAlaArgAspSerAlaAla.....	128
319	GAGAAGCCAGCGGTATCCAGTCTTCGCTGGAAGGGGAAACACTCCC	368
129ValPheAlaTrpLysGlyGluThrLeuG1	138
369	AGAATACTGGAGAAACACATACCGGCTCTCATCGCCAGATGGTCAAG	418
138	nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG	155
419	GGCCACACGAGTTCGATGATGGTGTGATGCTACACTCCTCATCTCC	468
155	lyProAspLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis	171
469	AAGGCTTCGAA.....TTCGAACACAGCCGGTGTCTTCCAGA	506
172	GluGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProAs	188
507	GCCACACAGAGCTGACACCTCGAATACCGCTCGGTCTTCTTCTACACTCA	556
188	pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle	205
557	AGCAGGTCTTCAACCAAGACAGAACCACTGGCACACAGTGTGTCGCGGC	606
205	ysGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg	221
607	ATGAACGGTGTTCGAGAGACAAACAGGTGTCCACCGCCTCTACCA	656
222	LeuValGlyValSerGluGluThrThrGlyValLysArgLeuTyrG1	238
657	GCTCGAGAGAGGAGGCAAACTCTCTTCCAGGCCATCAACGTCAACAGC	706
238	nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsps	255

```

707 CTGTTACAAAGTCCAAAGTTCCGATACATCTACGGCTGCGGCACTCCCTT 756
255 erValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeu 271
757 ATCGATGTTATCAACCGTCTCCGATGTCTATGATCGGGCAAGACAGC 806
272 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAl 288
807 TCTCGTCATGGTTAGCGGATGTGGCAAGGGTGGCGTCAATCCCTCC 856
288 aValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetL 305
857 GTGGCAAGCGCTCGCTTATCATCACAGAACTCGACCCCAATCTCGCT 906
305 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 321
907 CTCAGGCTGCCATGGAAGCTACCGATCGCGCGCATCGAGGAAGTCGT 956
322 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 338
957 CAAGGATGTCGATATCTCGTTACATGCACAGGAACTGGGATATCATCT 1006
338 lserGluAlaAspIlePheValThrThrGlyAsnLysAspIleIleM 355
1007 CTGTTGACATGATGGCCAGATGAAGGATATTGTGCGTAACATC 1056
355 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 371
1057 GGCCTACTCGATAAGCAATGATACAGATGGCGTCAATGAATACCCAGG 1106
372 GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrProG 388
1107 CATCAAGCACATCCCAATCAAGCAATACGACATGTGGGAATCCCG 1156
388 yValLysArgIleThrIleLysProGlnThrAspArgTyrPheProG 405
1157 ATGGCCACGCT...ATCCCTCTTCTGTGAGGGCGGCTTCTTAACCTT 1203
405 luthrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 421
1204 GGCTCGCTACAGGTCACCCATCTTTCGTTATGTCAATGTCATTCACAA 1253
422 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 438
1254 CCAGACACTCGCTACGCTCGACCTCTACGAAAGAGA.....GGAAATC 1297
438 nGlnValIleAlaGlnLeuGluLeuThrPasnGlnLysAlaSerGlyLysT 455
1298 TCGAGAGAAGGTTTACACACTTCGGAAGCATCTCGATGAAGAAGTCGCT 1347
455 yrGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAla 471
1348 CGCTCCACCTCGGATCTCGATGTCCACCTTACAAAGCTTACAGAA 1397
472 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAs 488
1398 CGAGGCTGACTACATCAAGTTCACAGTTCAGGTCCTTCAAGTCTGATG 1447
488 pGlnSerAspTyrValSerIleProIleGluGlyProTyrLysProProH 505
1448 CTTACCGTTAT 1458
505 istYrArgTyr 508

```

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:AA26500

seq_documentation_block:

ID AAR26500 standard; Protein: 486 AA.

XX AAR26500;

AC AAR26500;

XX 10-MAR-1993 (first entry)

DT

XX

Prod. of Nicotiana tabacum gene expressing at floral differentiation.
Flower; induction.

Nicotiana tabacum.

JP04258292-A.

14-SEP-1992.

14-FEB-1991; 91JP-0020702.

14-FEB-1991; 91JP-0020702.

(NISB) JAPAN TOBACCO INC.

WPI: 1992-354683/43.

N-PSDB; AAQ29419.

Gene expressing at floral differentiation for flowering control -
obtd. from culture of floral axis epithelium cells of Nicotiana
tabacum, for introduction into other plants

Disclosure; Page 6; 8pp; Japanese.

The protein sequence was deduced from the DNA sequence of a clone
obtd. by screening a cultured epithelial axis cDNA library with probes
from cultured and untreated mRNA. The clone obtd. expressed its
protein only at floral differentiation. The gene can be introduced
into other plants or can be suppressed by an antisense technique for
the control of flowering of plants.

Sequence 486 AA;

alignment_scores:

Quality: 1526.00 Length: 482

Ratio: 3.953 Gaps: 4

Percent Similarity: 80.083 Percent Identity: 63.071

alignment_block:

US-09-759-990-1 x AAR26500 ..

Align seg 1/1 to: AAR26500 from: 1 to: 486

34 GAGTACAGAAATGCCGACATCAACCTCCATCTCTCGGCGCTTAAGGAAC 83

14 GluTyrLysValLysAspMetSerGlnAlaAspPheGlyArgLeuGluI 30

84 TACCCCTTGTGAGAAGAAATGCCAGCTCTTATGTTCTTCGTGAGCGTT 133

30 eGluLeuAlaGluValGluMetProGlyLeuMetAlaCysArgThrGluP 47

134 ATTCGCTTCTAGCCATTCAGGGTGTCAAGGTGTCAAGATCTCTGGTCCCTCCAC 183

47 heGlyProSerGlnProPheLysGlyAlaLysIleThrGlySerLeuHis 63

184 ATGACAGTCCAGACAGCGCTCTCATCGAGACTCAGACTCAGCTCTTGGTGC 233

64 MetThrIleGlnThrAlaValLeuIleGluThrLeuThrAlaLeuGlyAl 80

234 TGATGTCAGATGGCTTCTGCAACATCTTCTTACACAGATACAGCG 283

80 aGluValArgTyrCysSerCysAsnIlePheSerThrGlnAspHisAlaA 97

284 CTGCTGCTATCTGTTGCGGCCCAACAGGCACACCAGAGAGCCAGCCGCT 333

97 laAlaAlaIleAlaArgAspSerAlaAla..... 106

334 ATCCAGTCTTCGCTGGAGGCGGAAACACTCCCGAATACTCGGAGAA 383

107ValPheAlaTyrLysGlyGluThrLeuGlnGluTyrTrpCp 121

384 CACATACCGCGCTCTCACATGCCAGATGGTCAAGCCCAAGCAGGTTG 433
121 sThrGluAlaLeuAspTrpGlyProGlyGlyProAspLeuilev 138
434 TCATGATGGTGGTGTACATCCCTCATCTCCAAAGGCTTC..... 477
138 alaAspGlyGlyAspAlaThrLeuLeuileHisGluGlyVallysAla 154
478GAATTCGAACAGCGGTGCTGTTCAGAGCCCAACAGACTGA 521
155 GluGluGluPheAlaLysAsnGlyThrileProAspProAsnSerThrAs 171
522 CAACCTCGAATACCGCTGGTCTGTGCTACACTCAAGCAGGCTTCAACC 571
171 pAsnAlaGluPheGlnLeuValLeuThrlelelelelelelelelele 188
572 AAGACAAGAACACCATGGCACAGTGTGTCGCGGCATGAACGGTGTTC 621
188 hrAspProLeuLysTyThrLysMetLysGluArgLeuValGlyValSer 204
622 GAAGAGACACAGAGGTCTCCACCGCTCTACAGCTCGAAGAGGAGG 671
205 GluGluThrThrGlyValLysArgLeuLysGlnMetGlnAlaAsnG 221
672 CAACCTCTCTCTCCAGCCATCAACGCTCAACGCTGTGTACAAAGTCCA 721
221 yThrLeuLeuPheProAlaIleAsnValAsnAspSerValThrLysSerL 238
722 AGTTCGATACATCTACGGCTGCGGCGCATCTCCCTTATCGATGTATCAAC 771
238 ysPheAspAsnLeuTyGlyCysArgHisSerLeuProAspGlyLeuMet 254
772 CGTGTCTCGATGTATGTCGCGGCAAGACAGCTCTCGTCATGGTTA 821
255 ArgAlaThrAspValMetIleAlaGlyLysValAlaLeuValAlaGlyTy 271
822 CGCGCATGTCCGCAAGGCTCGCTCAATCCCTCCGTGCGCAAGGCGCTC 871
271 rGlyAspValGlyLysGlyCysAlaAlaLeuLysGlnAlaGlyAlaA 288
872 GGTATATCATCAGAACATCGACCCATCTCGCTCTCCAGGCTGCCATG 921
288 rgValileValThrGluIleAspProleleCysAlaLeuGlnAlaThrMet 304
922 GAAGGCTACAGGTCGCGCGCATCGAGGAGTGTCTCAAGGATGTCTGATAT 971
305 GluGlyLeuGlnValLeuThrLeuGluAspValValSerAspValAspI 321
972 CTTCGTTTACATGACAGGAAGTGTGCGATATCATCTCTGTTCACATGATG 1021
321 ePheValThrThrGlyAsnLysAspIleIleMetValAspHisMetA 338
1022 CCAGATGAAGGATAAGGCTATTGTGCTTAACATCGGCGCTCTCGATAC 1071
338 rgLysMetLysAsnAsnAlaIleValCysAsnIleGlyHisPheAspAsn 354
1072 GAAATTCATACAGATGGCTCTCATGAATCCAGGATCAACAGCATCC 1121
355 GluIleAspMetLeuGlyLeuGluThrTyProGlyValLysLysIleTh 371
1122 AATCAAGCAGATACCATGATGTGGGAATTCACAGATGGCCAGCT...A 1168
371 rIleLysProGlnThrAspArgTrpValPheProAspThrAsnSerGlyI 388
1169 TCCTCTCTCTGTGAGGCGCGCTCTTAACTTGGCTGGCTACAGGT 1218
388 leileValLeuAlaGluGlyArgLeuMetAsnLeuGlyCysAlaThrGly 404
1219 CACCATCTCTGTGTATGCAATGTCTATCAAAACAGACACTCGCTCA 1268
405 HisProSerPheValMetSerCysSerPheThrAsnGlnValIleAlaG 421
1269 GCTCGACCTCTACCAAAAGAGA.....GGAAATCTCGAGAAGAAGGTTT 1312

421 nLeuGluLeuTripAsnGluLysSerSerGlyLysTyThrGluLysValT 438
1313 ACACACTTCGAGCATCTCGATGAAGAAGTCTCGCTCCACCTCGGA 1362
438 yrValLeuProProHisLeuAspGluLysValAlaAlaLeuHisLeuGly 454
1363 TCTCTCGATGTCCACCTTACAAAGCTTACACAGAGCAGGCTGACTACAT 1412
455 LysLeuGlyAlaLysLeuThrLysLeuSerLysAspGlnAlaAspTyrl 471
1413 CAACGTTCCAGTTCAGGCTCTCTCAAGCTCTGATGCTTACCGTTAT 1458
471 eSerValProValGluGlyProTyLysProAlaHisTyrlArgTyrl 486
seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAG34060
seq_documentation_block:
ID AAG34060 standard; Protein; 485 AA.
XX
AC AAG34060;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 41387.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

XX
PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 1523.00      Length: 487
  Ratio: 3.875          Gaps: 4
  Percent Similarity: 80.698      Percent Identity: 61.807

alignment_block:
US-09-759-990-1 x AAG34059

Align seg 1/1 to: AAG34059 from: 1 to: 510

19 ACTGGTCTCCATTCGAGTACAGAAATGCCGACATCAACCTCCATGTTCT 68
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
33 ThrSerSerGlyArgGluTyrLysValLysAspLeuSerGlnAlaAspPh 49
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
69 CGGCGGTAAAGAACTTACCCTTGTCTGAGAGAAATGCCAGGTCTTATGG 118
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
49 eGlyArgLeuGluIleGlnLeuAlaGluValGluMetProGlyLeuMeta 66
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
119 TTCTTCTGAGCGCTTATTCGCTTCTTAAGCCATTCAAGGGGTGTCAGAATC 168
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
66 laCysArgAlaGluPheGlyProSerLysProPheAlaGlyAlaArgIle 82
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
169 TCTGGTTCCTCCACATGACAGTCCAGACCGCTCCATCGAGACACT 218
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
83 SerGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 99
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
219 CACAGCTCTTGGTCTGATGTCAGATGGCTTCTCGCAACATCTTCTCTA 268
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
99 uThrAlaLeuGlyAlaGluValArgTyrPcysSerCysAsnIlePheSerT 116
   |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
269 CACAAGATACAGCGCTGCTGCTATCTGTTGTCGGCCCAACAGGCACCA 318

```

```

407 luThrAsnThrGlyIleLeValLeuAlaGlugLyArgLeuMetAsnLeu 423
1204 GGTCGCCTCAGGTACCCACTTTCGTTCATGTCAAATGCTCATTCACAAA 1253
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
424 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAS 440
1254 CGAGACACTCGCTCAGCTCGACCTCTACGAAAGAAGAGA.....GGAAATC 1297
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
440 nGlnValIleAlaGlnLeuGluLeuLutrrPlysGluLysSerSerGlyLysT 457
1298 TCAGAGAGAAGGTTTACACACTTCGGAAGCATCTCGATGAGAACCTCCCT 1347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457 YrGluLysLysValTyrrValLeuProLysHisLeuaspgLuLysValAla 473
1348 CGCCTCCACTCGGATCTTCGATGCCACCTTACAAGACTTACACAGAA 1397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
474 AlaLeuHisLeuGlyLysLeuGlyAlaLysLeuThrLysLeuThrLysSe 490
1398 GAGGCGTGACTACATCAACGTTCCAGTTGAGGGTCCCTTACAAGTCTGATG 1447
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
490 rGlnAlaaspTyrrIleSerValProIleGluGlyprotyrLysProAlah 507
1448 CTTACCCTTTAF 1458
|||||:|||||:
507 IsTyrrArgTyrr 510

seq_name: /SID55/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: AAG40087
seq_documentation_block:
ID   AAG40087 standard; Protein; 467 AA.
XX
AC   AAG40087;
DT
DE   18-OCT-2000 (first entry)
XX
KW   Arabidopsis thaliana protein fragment SEQ ID NO: 49692.
KW
KW   Protein identification; signal transduction pathway; metabolic
KW   hybridisation assay; genetic mapping; gene expression control;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
PN   EF1033405-A2.
XX
XX
XX   06-SEP-2000.
XX
XX   25-FEB-2000; 2000EP-0301439.
XX
XX   25-FEB-1999; 99US-0121825.
XX   05-MAR-1999; 99US-0123180.
XX   09-MAR-1999; 99US-0123548.
XX   23-MAR-1999; 99US-0125788.
XX   25-MAR-1999; 99US-0126264.
XX   29-MAR-1999; 99US-0126785.
XX   01-APR-1999; 99US-0127462.
XX   08-APR-1999; 99US-0128234.
XX   16-APR-1999; 99US-0128714.
XX   21-APR-1999; 99US-0129845.
XX   19-APR-1999; 99US-0130077.
XX   23-APR-1999; 99US-0130449.
XX   23-APR-1999; 99US-0130510.
XX   28-APR-1999; 99US-0130891.
XX   30-APR-1999; 99US-0131449.
XX   30-APR-1999; 99US-0132048.
XX   04-MAY-1999; 99US-0132407.
XX   05-MAY-1999; 99US-0132484.
XX   06-MAY-1999; 99US-0132485.
XX   06-MAY-1999; 99US-0132486.
XX   07-MAY-1999; 99US-0132487.
XX   11-MAY-1999; 99US-0132863.
XX   14-MAY-1999; 99US-0134256.
XX   14-MAY-1999; 99US-0134218.

```

```
PR 25-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1519.00      Length: 470
    Ratio: 3.966          Gaps: 4
Percent Similarity: 81.489 Percent Identity: 63.404

alignment_block:
US-09-759-990-1 x AAG40087 ..
Align seg 1/1 to: AAG40087 from: 1 to: 467

70 GCGCGTAAGGAACCTTACCCCTTGCTGAGAGAAATGCCAGGTCTTATGTT 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 GlyArgLeuGluLeuGluLeuAlaGluValGluMetProGlyLeuMetAl 23
120 TCTTGGTGAAGCGTTATCCCGTTCTAAGCCATTGAAGGGTGTCAATCT 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23 aCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgIle 40
170 CTGGTTCCTCCACATGACAGCCGCTCCTCATCGAGACACTC 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 hrgLysLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLeu 56
220 ACAGCTCTTGTCGATGATGAGTGGGCTTCTGCAACATCTTCTCTAC 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 ThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerTh 73
270 ACAAGATACAGCCGCTGCTATCGTTCGCGCCCAACAGGACACACAG 319
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 rGlnAspHisAlaAlaAlaAlaIleAlaArgaspSerAlaAla..... 87
320 AGAAGCCAGCCGGTATCCAGTCTTCGCTGGAAGGCGGAAACACTCCCA 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 .....ValPheAlaTrpLysGlyGluThrLeuGln 97
370 GAATACCTGGAGACACATACCGCTCTCACATGGCCAGATGGTCAAGG 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GluTyTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyL 114
420 CCCACACAGAGTTGTCGATGATGGTGTGATGTACACTCTCATCTCCA 469
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 yProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 131
470 AGGGCTTCGAA.....TTGAAACAGCCGCTGTCTCCAGAG 507
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
131 LuGlyValLysAlaGluGluIlePheGluLysThrGlyGlnValProasp 147
508 CCAACAGAGAGTCAACACCTCGAATACCGCTGCTTCTTGTACACTCAA 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle 164
558 GCAGGTCTTCAACCAAGAACACACTGGCAGACAGTGTCTGCCGGCA 607
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
164 sGluGlyLeuGlnValAspProLysLysTyHisLysMetLysGluArgL 181
608 TGAACGGTGTTCGAGAGACACACAGGTGTCCACCGCTCTTACAG 657
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
181 euValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyGln 197
658 CTCGAGAGGAGGAGCAAACTCTCTTCCAGCCATCAACGTCAACGAGCG 707
||| ||| ||||| ||||| ||||| ||||| ||||| |||||
198 MetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAspSe 214
```

708 TGTTACAAAGTCCAAAGTTCGATACATCTAGCGCTGGCGCCACTCCCTTA 757
214 rValThrLysSerLysPheAspAsnLeuTyrGlyCysArgHisSerLeuP 231
758 TCGATGGTATCAACCGTTCGATGTCATGATCGGCGGCAAGACAGCT 807
231 roAspGlyLeuMetArgAlaThrAspValMetLeuAlaGlyLysValAla 247
808 CTCGTCATGGTTCACGGCGATGCGGAAGGCTGCGCTCAATCCCTCCG 857
248 ValIleCysGlyTyrGlyAspValGlyLysGlyCysAlaAlaMetLy 264
858 TGGCCCAAGCGCTCGGCTTATCATCACAGTCCGAGCTCGAGGAGTCTGC 907
264 sthrAlaGlyAlaArgValIleValThrGluLeuAspProIleCysAla 281
908 TCCAGGCTGCCATGGAAGCTACAGGCTCGCGCATCGAGGAGTCTGC 957
281 euGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVal 297
958 AAGGATGTCGATATCTTCGTTACATGACAGGAACTGCGATATCATCTC 1007
298 SerGluAlaAspIlePheValThrThrThrGlyAsnLysAspIleIleMe 314
1008 TGTTCACATGATGGCGCATGAAGCTAAGGCTATTGTCGGTAAACATCG 1057
314 tValAspHisMetArgLysMetLysAsnAsnAlaIleValIcysAsnIleG 331
1058 GCCACTTCGATCAACGAATGTATACAGATGGCTCATGAATACCCAGGC 1107
331 lyHisPheAspAsnGluLeuAspMetLeuGlyLeuGluThrTyrProGly 347
1108 ATCAAGCACATCCCAATCAAGCAGACATACGATGTTGGAAATCCGAGA 1157
348 ValLysArgIleThrIleLysProGlnThrAspArgTyrValPheProG 364
1158 TGGCCACGCT...ATCCRCCTCTTGTGAGGCGCCTTCTTAACCTTG 1204
364 uThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeuG 381
1205 GCTCGGCTACAGGTCACCCATCTTCGTTATATCATGTCATTCACAAAC 1254
381 lyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAsn 397
1255 CAGACACTCGCTCAGCTCGACTCTACGAAAGAGA.....GGAATCT 1298
398 GlnValIleAlaGlnLeuGluLeuThrAsnGluLysAlaSerGlyLysTy 414
1299 CGAGAAGAGTTTACACTTCCGAGCATCTCGATGAAGAAGTGGCTC 1348
414 rGluLysLysValTyrValLeuProLysHisLeuAspGluLysValAlaL 431
1349 GCCTCCACCTCGGATCTCGATGTCACCTTACAAAGTTACAGAGAAG 1398
431 euLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLysAsp 447
1399 CAGGCTGACTACATCAAGTTCAGGTCAGGTCCTTACAAGTCTGATCC 1448
448 GlnSerAspTyrValSerIleProIleGluGlyProTyrLysProProHi 464
1449 TTACCGTTAT 1458
464 stYrArgTyr 467

seq_name: /SID85/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA41407

seq_documentation_block:

ID: AA41407 standard; Protein: 485 AA.

XX AC AA41407;

XX DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 51514.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 18-MAY-1999; 99US-0134370.

XX 19-MAY-1999; 99US-0134768.

XX 20-MAY-1999; 99US-0134941.

XX 21-MAY-1999; 99US-0135124.

XX 24-MAY-1999; 99US-0135353.

XX 25-MAY-1999; 99US-0135629.

XX 27-MAY-1999; 99US-0136021.

XX 28-MAY-1999; 99US-0136392.

XX 01-JUN-1999; 99US-0136782.

XX 03-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 14-JUN-1999; 99US-0138847.

XX 16-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 17-JUN-1999; 99US-0139453.

XX 17-JUN-1999; 99US-0139492.

XX 18-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

XX 18-JUN-1999; 99US-0139456.

XX 18-JUN-1999; 99US-0139457.

XX 18-JUN-1999; 99US-0139458.

XX 18-JUN-1999; 99US-0139459.

XX 18-JUN-1999; 99US-0139460.

XX 18-JUN-1999; 99US-0139461.

XX 18-JUN-1999; 99US-0139462.

XX 18-JUN-1999; 99US-0139463.

XX 18-JUN-1999; 99US-0139750.

XX 18-JUN-1999; 99US-0139763.

XX 21-JUN-1999; 99US-0139817.

us-09-759-990-1.n2p.rag

Mon Aug 19 15:16:18 2002

```

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141387.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-014342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 1512.00      Length: 487
Ratio: 3.887          Gaps: 4
Percent Similarity: 79.877 Percent Identity: 61.396

alignment_block:
US-09-759-990-1 x AAG41407 ..
Align seg 1/1 to: AAG41407 from: 1 to: 485
19 ACTGGTGTCCATTTCGAGTACAGAAATTCGCGACATCAACCTCCATGTTCT 68
8 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 24
69 CGCGCGTAAGGAACCTTACCCTTCGTCGAGAGGAATGCCAGGTCCTTATGG 118
24 eGlyArgLeuGluIleGluLeuAlaGluValGluMetProGlyLeuVal 41
119 TTCTTCGTGAGCGTWTATCCGTTCTTAAGCCATTGAAGGTCGTCAGATC 168
41 exCysValThrGluPheGlyProSerGlnProLeuLysGlyAlaArgile 57
169 TCTGGTTCCTCCACATGACAGTCCAGACAGCGCTCCTCATCGACACT 218
:::|||||

```


Mon Aug 19 15:16:18 2002

PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	99US-01324863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0132456.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	03-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137328.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139463.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140991.	PR	15-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0142154.	PR	22-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-0155486.
PR	06-JUL-1999;	99US-0142390.	PR	24-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142920.	PR	29-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	06-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144085.	PR	08-OCT-1999;	99US-0158232.
PR	16-JUL-1999;	99US-0144086.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.

```

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 1512.00
  Ratio: 3.887
  Percent Similarity: 79.877
  Gaps: 4
  Percent Identity: 61.396

alignment_block:
  US-09-759-990-1 x AAG41406 ..
  Align seg 1/1 to: AAG41406 from: 1 to: 497

19 ACTGGTCTCCATTCGAGTACAGAAATGCGGACATCAACCTCCATGTTCT 68
20 ThrSerSerGlyArgGluThrLysValLysAspMetSerGlnAlaAspPh 36
69 CGGCGGTAAAGAACTTACCTTCGTCGAGAGAAATGCCAGGTCTTATGG 118
36 eGlyArgLeuGluLeuGluLeuValGluMetProGlyLeuValS 53
119 TTCTCTGAGCGTTATPCCGCTTCTAAGCCATTGAAGGTGTTCAGAAATC 168
53 erCysValThrGluPheGlyProSerGlnProLeuLysGlyAlaArgIle 69
169 TCTGGTTCCTCCACATGACAGTCCAGACAGCCGCTCTCATCGACACT 218
70 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuLeuGluThrLe 86
219 CACAGCTCTTTGGTCTGATGTCAGATGGCTTCCTGCAACACTTCTCTA 268
86 uThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerT 103
269 CACAAGATACAGCGCTGCTGCTATCGTTCGTTCGGCCCAACAGGCACCA 318
103 hrGlnAspHisAlaAlaAlaAlaIleAlaAa-gaspSerAlaAla..... 117
319 GAGAAGCCAGCGGTATCCCGATCTTCGCTTCGCGGAGGCGGAAACACTCC 368
118 .....ValPheAlaTrpLysGlyGluThrLeuG 127
369 AGAATACTGGGAAACACATACCGCTCTCATGTCGACATGGCCAGTCAAG 418
127 nGluTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGly 144
419 GCCACAGCAGGTTGTCGATGGTGGTGTGATGTCACACTCCTCATCTCC 468
144 lypProSerLeuIleValAspGlyGlyAspAlaThrLeuLeuIleHis 160
469 AAGGGTTCGAA.....TTCCGAAACAGCCGGTGTGTTCGACGA 506
161 GluGlyValLysAlaGluGluIlePheAlaLysAsnGlyThrPheProAs 177
507 GCCAAGAGAGTGAACACTCGAATACCGCTGGCTTCTGCTACACTCA 556
177 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 194
557 AGCAGGTCTTCAACCAAGACAAGAACCACTGGGCACACAGTGTGTCGGGC 606

```

```

119 ySaspGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 210
607 ATGAACGGTGTTCGAGAGACAAACAGGTGTCACCGCTCTACCA 656
211 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTrpG 227
657 GCTCGAAGAGGAGGCAACTCTCTTCCAGCCATCAACGTCACAGCG 706
227 nMetGlnGluThrGlyAlaLeuLeuPheProAlaIleAsnValAsnAsp 244
707 CTGTTACAAAGTCCAAAGTTCGATAACATCTAGCGCTCGCGCCACTCC 756
244 erValThrLysSerLysPheAspAsnLeuTrpGlyCysArgHisSerLeu 260
757 ATCGATGGTATCAACCGCTTCGATGTCATGATCGCGCGCAAGACAGC 806
261 ProAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysVal 277
807 TCTCGTCAATGGTTACGGCATGTCGCAAGGCTGCGCTCAATCCCTCC 856
277 aValIleCysGlyTrpGlyAspValGlyLysGlyCysAlaAlaMetL 294
857 GTGGCAAGCGCTCGCGTTATCATCACAGAACTCGACCAATCTCGCT 906
294 ysThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAla 310
907 CTCACGGCTGCCATGGAAGGCTACCGGTCGCGCATCGAGGAAGTCGT 956
311 LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspVal 327
957 CAAGATGTCGATATCTCTTACATGCACAGAACTGCGATATCATCT 1006
327 lSerGluAlaAspIlePheCysThrThrGlyAsnLysAspIleIleM 344
1007 CTGTTGACATGATGGCCCATGAGGATAGGCTATTGTCGGTAAACATC 1056
344 etValAspHisMetArgLysMetLysAsnAlaIleValCysAsnIle 360
1057 GCCCACTTCGATAACGAAATGATACAGATGGCTCATGAATACCCAGG 1106
361 GlyHisPheAspAsnGluLeuAspMetLeuGlyLeuGluThrTrpPro 377
1107 CATCAAGCACATCCCAATCAAGCAGAAATACGACATGTGGAAATCCAG 1156
377 yValLysArgIleThrIleLysProGlnThrAspArgTrpValPhePro 394
1157 ATGGCCACGCT...ATCCTCCTCTTGTGAGGCGCGCTTCTTAACTT 1203
394 spThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 410
1204 GGCTCGCTACAGGTCAACCATCTTTCTGTTATGTCATGTCATTCACAA 1253
411 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThr 427
1254 CCAGACACTCGCTACGCTGAGCTCTTACGAAAGAGA.....GGAATC 1297
427 nGlnValIleAlaGlnLeuGluLeuTrpAsnGluLysSerSerGlyLys 444
1298 TCAGAAGAAGGTTTACACACTTCGAGCATCTCGATGAGGAAGTCGCT 1347
444 yrGluLysLysValTrpValLeuProLysHisLeuAspGluLysValAla 460
1348 CGCTCCACCTCGGATCTCTGATCTCCACTTACAAAGCTTACACAGAA 1397
461 AlaLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLys 477
1398 GCAGGCTCACTACATCAAGTTCAGTTCAGGCTTCTTACAAGTCTGATG 1447
477 pGlnSerAspTrpValSerIleProValGluGlyProTyrLysProValH 494
1448 CTTACCGTTAT 1458

```



```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 1494.00
    Ratio: 3.952
    Percent Similarity: 80.426
    Percent Identity: 62.766

alignment_block:
US-09-759-990-1 x AAG41408 ..

Align seg 1/1 to: AAG41408 from: 1 to: 467

70 GGCGTAAAGAACTTACCTTGTGAGAAAGAAATGCCAGGTCTTATGGT 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7 GlyArgLeuGlutIleGluLeuAlaGluValGluMetProGlyLeuValSe 23
120 TCTTCGTGAGCGTTATTCGGCTTCTAAGCCATTGAAGGTGTCAAACT 169
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 rCysValThrGluPheGlyProSerGlnProLeuLysGlyAlaArgIle 40
170 CTGGTTCCTCCACATCAGTCAGCCAGCGCTCTCATCGAGACACATC 219
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 hrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLeu 56
220 ACAGCTCTTGGTGTGATGTCAGATGGCTTCTCTCAACATCTTCTCTAC 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
57 ThrAlaLeuGlyAlaGluValArgTrpCysSerCysAsnIlePheSerTh 73
270 ACAAGATACAGCCCTGCTGCTATCGTTGTGCGCCCAACAGGCACACCA 319
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
73 rGlnAspHisAlaAlaAlaAlaIleAlaArgAspSerAlaAla..... 87
320 AGAAGCCAGCGGTATCCAGTCTTCGCTGTGAGGGCGGAAACACTCCCA 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 .....ValPheAlaTrpLysGlyGluThrLeuGln 97
370 GAATACTGGGAGAACATACCGGCTCTCATGCGCCAGATGGTCAAGG 419
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GlutyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 114
420 CCCACAGCAGGTTCTCGATGATGGTGTGATGCTACATCTCTCATCTCCA 469
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 yProAspIleValAspGlyGlyAspAlaThrLeuLeuIleHisG 131
470 AGGCTTCTCGAA.....TTCGAAACAGCGGTGCTGTCTCCAGAG 507
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 luGlyValLysAlaGluGluIlePheAlaLysAsnGlyThrPheProAsp 147
508 CCAACAGAGCTGACACCTCGAATACCGCTCGGTCTTGTCTACACTCAA 557
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 ProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIle 164
558 GCAGGTCTTCAACCAAGAACACACCTGGCACACAGTTGCTGCCGGCA 607
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
164 sAspGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 181
608 TGAACGGTGTTCGAGAGACAAACACAGGTGTCCACCGCTCTTACCAG 657

```

```

181 euValGlyValserGluThrThrThrGlyValLysArgLeuTyrgln 197
    :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 CTCGAGAAGGAGGGCAAACTCCTCTTCCAGCCATCAACGTCAACGAGCG 707
    : : : : : ||| ||||| ||||| ||||| ||||| ||||| |||||
198 MetGlnGluThrGlyAlaLeuLeuPheProAlaIleasnValasnAspSe 214
    TGTGTACAAAGTCCAGTTCGATAACATCTACGGCTCGCGCCACTCCCTTA 757
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
214 rValThrLysSerLysPheAspAsnLeuTyrgLysArgHisSerLeuP 231
    TCGATGTATCAACCGTGTCCGATGTCTATCATCGCGCGCAACAGACGT 807
    ||||| :: ||||| ||||| ||||| ||||| ||||| ||||| |||||
231 roAspGlyLeuMetArgAlaThrAspValMetIleAlaGlyLysValAla 247
    808 CTGCTCATGGTTACGGCGATGTTCGGCAAGGGCTGGCTCAATCCCTCCG 857
    : : : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
248 ValIleCysGlyTyrgLysPValGlyLysGlyCysAlaAlaAlaMetLy 264
    TGCCCAAGGGGCTCGCGTTATCATCACAGAATCGACCAANTCTGGCGTC 907
    sThrAlaGlyAlaArgValIleValThrGluIleAspProIleCysAlaL 281
    908 TCCAGGCTGCCATGGAGGCTACACAGTCCGCGCGCGCATCGAGGAAGTCGTC 957
    euGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVal 297
    958 AAGGATGTCTGATATCTCGTTACATGCACAGAATCGAGAACTGCGATCATCTC 1007
    : : : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 SerGluAlaAspIlePheCysThrThrThrGlyAsnLysAspIleIleMe 314
    TGTGTGACATGATGGCCAGATGAAGGATAAGGCTATTGTTCGTTAACTCG 1057
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
314 tValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIleG 331
    GCCACTTCGATAACGAATTCATACAGATGCGCTCATGAATACCCAGGC 1107
    331 LysHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrgLy 347
    1108 ATCAGCACATCCCAATCAAGCCAGAATACGACATGTGGGAATTCGCCAGA 1157
    : : : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
348 ValLysArgIleThrIleLysProGluThrAspArgTyrPValPheProAs 364
    1158 TGCCCAACGCT...ATCTCTCTTCTGTGAGGGCGGCTTCTTAACCTTG 1204
    pThrAsnSerGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeuG 381
    1205 GTCGCTACAGGTCACCCATCTTTCTTATGTCAATCTCATTCATCAACAAC 1254
    381 LysCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAsn 397
    1255 CAGACATCGCTCAGCTGCAGCTCTACGAAAGAGA.....GGAAATCT 1298
    GlnValIleAlaGlnLeuGluLeuTyrPasnGluLysSerSerGlyLysTy 414
    1299 CGAGAAGAAGGTTTACACACTTCCGAAGCATCTCGATCAAGAAGTCGCTC 1348
    rGluLysLysValTyrvAlaLeuProLysHisLeuAspGluLysValAlaLa 431
    1349 GCCTCCACTCGGATCTCTCGATGTCACCTTACAAAGCTTTACACAGAG 1398
    431 laLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuThrLysAsp 447
    CAGGCTGACTACATCAACGCTCCAGTTGAGGTCCTTACAGAGTCTGATGC 1448
    GlnSerAspTyrvAlaSerIleProValGluGlyProIlyrLysProValLH1 464
    1449 TTACCGTTAT 1458
    ||||| |||||
464 slyrArgTyr 467

```

seq_documentation_block:	
ID	AAW01458 standard; Protein; 485 AA.
XX	
AC	AAW01458;
XX	
DT	22-FEB-1997 (first entry)
XX	
DE	Asparagus S-adenosyl-L-homocysteine hydrolase.
XX	
KW	S-adenosyl-L-homocysteine hydrolase; SHH; promoter; monocot; dicot;
KW	transgenic plant; disease resistance; pathogen resistance.
XX	
OS	Asparagus officinalis.
XX	
FH	Key Location/Qualifiers
FT	Region 150..190
FT	/note= "region found in SHH of photosynthetic
FT	species"
FT	Binding-site 263..294
FT	/label= NAD+_binding_site
XX	
PN	WO9632488-A1.
XX	
PD	17-OCT-1996.
XX	
PF	10-APR-1996; 96WO-GB00882.
XX	
PR	10-APR-1995; 95GB-0007381.
XX	
PA	(ZENE) ZENECA LTD.
XX	
PI	Draper J, Greenland AJ, Skipsey M, Warner S;
XX	
DR	WPI; 1996-477138/47.
DR	N-PSDB; AAT44513.
XX	
PT	S-adenosyl-L-homocysteine hydrolase promoter - used for driving
PT	expression of effector genes, such as pathogen resistance genes, in
PT	transgenic plants
XX	
PS	Example 1; Fig 1; 57pp; English.
XX	
CC	Asparagus S-adenosyl-L-homocysteine hydrolase (SHH) (AAW01458)
CC	catalyses the reversible hydrolysis of S-adenosyl-L-homocysteine
CC	to adenosine and homocysteine. It contains an extra stretch of
CC	amino acid residues (positions 150-190) previously found in other
CC	photosynthetic species, parsley and Rhodobacter capsulatus, but not
CC	in SHHs from non-photosynthetic species. A cDNA sequence (AAT44513)
CC	coding for the asparagus was used to identify the Arabidopsis
CC	thaliana SHH gene (AAT44515) and promoter (AAT44514), useful for
CC	expression of effector genes in transgenic plants.
XX	
SQ	Sequence 485 AA;
XX	

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:AAW01458

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 1473.00      Length: 459
  Ratio: 3.939          Gaps: 4
  Percent Similarity: 81.481  Percent Identity: 63.181

alignment_block:
US-09-759-990-1 x AAG34061 ..

Align seg 1/1 to: AAG34061 from: 1 to: 450

103 ATGCCAGGCTTATGCTGCTGAGCGGTATTCCGCTTCTAAGCCATT 152
      |||||
1 MetProGlyLeuMetAlaCysArgAlaGluPheGlyProSerLysProPh 17
      |||||

153 GAAGGGTGCAGAACTCTGTTCCCTCCACATGACAGTCCAGCAGCG 202
      : |||||
17 eAlaGlyAlaArgIleSerGlySerLeuHisMetThrIleGlnThrAlav 34
      : |||||

203 TCCTCATCGAGACATCACAGCTCTTGGTGTGATGTCAGATGGCTTCC 252
      |||||
34 albuileGluThrLeuThrAlaLeuGlyAlaGluValArgTrpCysSer 50
      |||||

253 TGCAACATCTTCTACAAAGATACAGCGCTGCTGCTATGCTGTGTCGG 302
      |||||
51 CysAsnIlePheSerThrGlnAspHisAlaAlaAlaIleAlaArgAs 67
      |||||

303 CCCAACAGGCACACAGAGCCAGCGGTATCCAGTCTTGGCTGGA 352
      : |||||
67 pSerAlaAla.....ValPheAlaIleTrpL 75

353 AGGGCGAAACACTCCAGATACTGGGAGAACACATACCGGCTCTCACA 402
      |||||
75 ysglyGluThrLeuGluGluTrpTrpCysThrGluArgCysLeuAsp 91
      |||||

403 TGCGCAGATGGTCAAGGCCACACAGCTTGTTCGATGATGGTGTATGC 452
      |||||
92 TrpGlyGluAlaGlyGlyProAspLeuIleValAspAspGlyGlyAspAl 108
      |||||

453 TACATCTCTCATCTCCAAAGGCTTC.....GAATTCGAAACAG 490
      |||||
108 aThrLeuLeuIleHisGlyValLysAlaGluGluTrpGluLysT 125
      |||||

491 CCGGTGTGTTCCAGAGCCACAGAGCTGACAACTCGAATACCGCTGC 540
      : |||||
125 hrGlyLysIleProAspProGluSerThrAspAsnAlaGluPheLysIle 141
      |||||

541 GTTCTTCTACATCAAGCGGTCTTCAACCAAGACAGCAACCACTGGCA 590
      |||||
142 ValLeuThrIleIleArgAspGlyLeuLysAlaAspProLysLysTrp 158
      : |||||
591 CACAGTGTGCTCGCGCATGAACGGTGTTCGAGAGACAAACAGAGTG 640
      : |||||
158 gLysMetLysGluArgLeuValGluValSerGluGluThrThrGly 175
```

```
641 TTCACGCGCTTACCAGCTCGAGAAGGAGGCGCAAACTCCTCTCCAGCC 690
      |||||
175 alLysArgLeuTyrGlnMetGlnGluThrGlyAlaLeuLeuPheProAla 191
      |||||

691 ATCAACGCTCAACGACGCTGTTCACAAAGTCCAAAGTTCGATACATCTACGG 740
      |||||
192 IleAsnValAsnAspSerValThrLysSerLysPheAspAsnLeuTyrGl 208
      |||||

741 CTGCGCGCTCCTCCCTTATCGATGATGATCAACCGCTTCCGATGTCATGA 790
      |||||
208 yCysArgHisSerLeuProAspGlyLeuMetArgAlaThrAspValMetI 225
      |||||
791 TCGCGCGCAAGACAGCTCTCTCATGGTTTACGGCGATGTCGGCAAGGCC 840
      |||||
225 leaLagLysValAlaValValCysGlyTyrGlyAspValGlyLysGly 241
      |||||

841 TGGCTCAATCCCTCCGTCGCGCAAGGCGCTCGCTTATCATCACAGAAGT 890
      |||||
242 CysAlaAlaAlaLeuLysGlnAlaGlyAlaArgValIleValThrGluI 258
      |||||

891 CGACCCCAATCTCGCTCTCCAGGCTGCGATGAGAGCTACCAAGGTCGCGC 940
      |||||
258 easPProIleCysAlaLeuGlnAlaLeuMetGluGlyLeuGlnValLeuP 275
      |||||

941 GCATCGAGGAAGTCGTCAAGGATCTCGATATCTTCTGTTTACATGACAGGA 990
      : |||||
275 roLeuGluAspValValSerGluAlaAspIlePheValThrThrGly 291
      : |||||

991 AACTGCGATATCATCTCTGTGACATGATGCGCCAGATGAAGGATAAGGC 1040
      |||||
292 AsnLysAspIleIleMetValAspHisMetArgLysMetLysAsnAla 308
      |||||

1041 TATTGTCGTAACATCGCCACTTCTCGATTAACGAAATTGATACAGATGGCC 1090
      |||||
308 alIeValCysAsnIleGlyHisPheAspAsnGluIleAspMetLeuGlyL 325
      |||||

1091 TCATGAATATCCAGGATCAAGCACATCCCAATCAAGCCAAATACGAC 1140
      |||||
325 euGluThrTyrProGlyValLysArgIleThrIleLysProGlnThrAsp 341
      |||||

1141 ATGTGGGAATTCGCCAGATGGCCAC...GCTATCTCTCTTCTGCTGAGGG 1187
      |||||
342 ArgTrpValPheProGluThrAsnThrGlyIleIleValLeuAlaGluI 358
      |||||

1188 CCGCTTCTTAACCTTGGCTGCGCTACAGGTACCCATCTTCTGTTATGT 1237
      |||||
358 yArgLeuMetAsnLeuGlyCysAlaThrGlyHisProSerPheValMet 375
      |||||

1238 CAATGTCATTCACAAACCCAGACATCGCTCAGCTCAGCTTACGAAAG 1287
      |||||
375 erCysSerPheThrAsnGlnValIleAlaGlnLeuGluLeuIleLysGlu 391
      |||||

1288 AGA.....GGAAATCTCGAGAAGAGTTTACACACTTCCGAAAGCATCT 1331
      : |||||
392 LysSerSerGlyLysTyrGluLysLysValTyrValLeuProLysHisLe 408
      |||||

1332 CGATGAAGAAGTCTGCTCGCTCCACTCCGATCTCGATGTCACCTTA 1381
      |||||
408 uAspGluLysValAlaAlaLeuHisLeuGlyLysLeuGlyAlaLysLeu 425
      |||||

1382 CAAAGCTTACACAGAGGAGGTGCTACTACATCAACGTTCCAGTTGAGGGT 1431
      |||||
425 hrLysLeuThrLysSerGlnAlaAspTyrIleSerValProIleGluGly 441
      |||||

1432 CTTTACAAGTCTGATCTTACCGTTAT 1458
      |||||
442 ProTyrLysProAlaHisTyrArgTyr 450
      |||||

seq.name: /SDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAG22957
seq_documentation_block:
ID AAG22957 standard; Protein; 467 AA.
```

XX AC AAG222957; 99US-0139462.
XX AC 17-OCT-2000 (first entry) 99US-0139463.
XX DT 99US-0139750.
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26083. 99US-0139763.
XX DE 99US-0139817.
XX DE 99US-0139899.
XX DE 99US-0140353.
XX DE 99US-0140354.
XX DE 99US-0140695.
XX DE 99US-0140823.
XX OS Arabidopsis thaliana. 99US-0140991.
XX XX EP1033405-A2. 99US-0141287.
XX PN 99US-0141842.
XX XX 99US-0142154.
XX PD 99US-0142055.
XX PF 99US-0142390.
XX XX 99US-0142803.
XX XX 99US-0142920.
XX XX 25-FEB-2000; 2000EP-0301439. 99US-0142977.
XX XX 99US-0143542.
XX XX 99US-0143624.
XX XX 99US-0144005.
XX XX 99US-0144085.
XX XX 99US-0144086.
XX XX 99US-0144325.
XX XX 99US-0144331.
XX XX 99US-0144332.
XX XX 99US-0144333.
XX XX 99US-0144334.
XX XX 99US-0144335.
XX XX 99US-0144352.
XX XX 99US-0144632.
XX XX 99US-0144884.
XX XX 99US-0144814.
XX XX 99US-0145086.
XX XX 99US-0145088.
XX XX 99US-0145085.
XX XX 99US-0145087.
XX XX 99US-0145089.
XX XX 99US-0145192.
XX XX 99US-0145145.
XX XX 99US-0145218.
XX XX 99US-0145224.
XX XX 99US-0145276.
XX XX 99US-0145913.
XX XX 99US-0145918.
XX XX 99US-0145919.
XX XX 99US-0145951.
XX XX 99US-0146386.
XX XX 99US-0146388.
XX XX 99US-0146389.
XX XX 99US-0147038.
XX XX 99US-0147204.
XX XX 99US-0147302.
XX XX 99US-0147192.
XX XX 99US-0147260.
XX XX 99US-0147303.
XX XX 99US-0147416.
XX XX 99US-0147493.
XX XX 99US-0147935.
XX XX 99US-0148171.
XX XX 99US-0148319.
XX XX 99US-0148341.
XX XX 99US-0148565.
XX XX 99US-0148684.
XX XX 99US-0149368.
XX XX 99US-0149175.
XX XX 99US-0149426.
XX XX 99US-0149722.
XX XX 99US-0149723.
XX XX 99US-0149929.
XX XX 99US-0149902.
XX XX 99US-0149930.

us-09-759-990-1.n2p.rag

```

25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
26-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 1457.00
  Ratio: 3.865
  Percent Similarity: 80.901
  Length: 466
  Gaps: 4
  Percent Identity: 62.232

alignment_block:
US-09-759-990-1 x AAG22957

Align seg 1/1 to: AAG22957 from: 1 to: 467

19 ACTGGTGCCTCCATTCCGATACAGAAATTCGCGACATCAACCTCCATGTTCT 68
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 ThrSerSerGlyArgGluTrpLysValLysAspMetSerGlnAlaAspPh 24
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 CGGCCGTAGGAACCTTACCCTTCCGTCGAGAGGAAATCCCGAGGCTTTAGG 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 eGlyArgLeuGluGluAlaGluValGluMetProGlyLeuMet 41
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 TTTCTGTCGAGCCCTTATCCCGCTTCTAAGCCATTGAAGGGTGTTCAGAAATC 168

```

```
332 etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysAsnIle 348
1057 GGCCTACTCGATAACGAATGTATACAGATGGCTCATGAATATACCCAGG 1106
|||||
349 GlyHisPheAsnGluIleAspMetLeuGlyLeuGluThrTyrProG1 365
|||||
1107 CATCAAGCACATCCCAATCAAGCAGATACGACATGTGGAAATCCAG 1156
|:::|
365 yValLysArgIleThrIleLysProGlnThrAspArgTyrValPheProG 382
|||||
1157 ATGGCACGCT...ATCCTCTCTTCTGCTGAGGCGCGCTTCTTAACCTT 1203
:: |||
382 luThrLysAlaGlyIleIleValLeuAlaGluGlyArgLeuMetAsnLeu 398
|||||
1204 GCGTGGCTACAGGTACCCATCTTCGTGTATGTCATGATTCACAA 1253
|||||
399 GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 415
|||||
1254 CCAGACATCGCTCAGCTCGACCTCTACGAAAGAGA.....GGAATC 1297
|||||
415 nGlnValIleAlaGlnLeuGluLeuIlePheAsnGluLysAlaSerGlyLysT 432
|||||
1298 TCGAAGAAGAGTTTACACATCTCCGAAGCATCTCGATGAAGAAGTCGCT 1347
|||||
432 yrGluLysLysValTyrValIleuProLysHisLeuAspGluLysValAla 448
|||||
1348 CCGCTCCACCTCGGATCTCTCGATGTCACCTTACAAAGTTTACACAG 1395
|||||
449 LeuLeuHisLeuGlyLysLeuGlyAlaArgLeuThrLysLeuSerLys 464

seq_name: /SID5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA22956
seq_documentation_block:
ID   AAG22956 standard; Protein: 489 AA.
XX
AC   AAG22956;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 26082.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999; 99US-0121825.
PR   05-MAR-1999; 99US-0123180.
PR   09-MAR-1999; 99US-0123548.
PR   23-MAR-1999; 99US-0125788.
PR   25-MAR-1999; 99US-0126284.
PR   29-MAR-1999; 99US-0126785.
PR   01-APR-1999; 99US-0127462.
PR   06-APR-1999; 99US-0128234.
PR   08-APR-1999; 99US-0128714.
PR   16-APR-1999; 99US-0129845.
PR   19-APR-1999; 99US-0130077.
PR   21-APR-1999; 99US-0130449.
PR   23-APR-1999; 99US-0130510.
PR   28-APR-1999; 99US-0130891.
PR   30-APR-1999; 99US-0131449.
PR   30-APR-1999; 99US-0132048.
PR   04-MAY-1999; 99US-0132407.
PR   05-MAY-1999; 99US-0132484.
PR   06-MAY-1999; 99US-0132485.
PR   06-MAY-1999; 99US-0132486.
PR   06-MAY-1999; 99US-0132487.
PR   07-MAY-1999; 99US-0132487.
PR   11-MAY-1999; 99US-0132486.
PR   14-MAY-1999; 99US-0134218.
PR   14-MAY-1999; 99US-0134219.
PR   14-MAY-1999; 99US-0134221.
PR   14-MAY-1999; 99US-0134370.
PR   18-MAY-1999; 99US-0134768.
PR   19-MAY-1999; 99US-0134941.
PR   20-MAY-1999; 99US-0135124.
PR   21-MAY-1999; 99US-0135353.
PR   25-MAY-1999; 99US-0135629.
PR   27-MAY-1999; 99US-0136392.
PR   28-MAY-1999; 99US-0136782.
PR   01-JUN-1999; 99US-0137222.
PR   03-JUN-1999; 99US-0137528.
PR   04-JUN-1999; 99US-0137502.
PR   07-JUN-1999; 99US-0137724.
PR   08-JUN-1999; 99US-0138094.
PR   10-JUN-1999; 99US-0138540.
PR   10-JUN-1999; 99US-0138847.
PR   14-JUN-1999; 99US-0139119.
PR   16-JUN-1999; 99US-0139452.
PR   16-JUN-1999; 99US-0139453.
PR   17-JUN-1999; 99US-0139492.
PR   18-JUN-1999; 99US-0139454.
PR   18-JUN-1999; 99US-0139455.
PR   18-JUN-1999; 99US-0139456.
PR   18-JUN-1999; 99US-0139457.
PR   18-JUN-1999; 99US-0139458.
PR   18-JUN-1999; 99US-0139459.
PR   18-JUN-1999; 99US-0139460.
PR   18-JUN-1999; 99US-0139461.
PR   18-JUN-1999; 99US-0139462.
PR   18-JUN-1999; 99US-0139463.
PR   18-JUN-1999; 99US-0139750.
PR   18-JUN-1999; 99US-0139763.
PR   21-JUN-1999; 99US-0139817.
PR   22-JUN-1999; 99US-0139899.
PR   23-JUN-1999; 99US-0140353.
PR   23-JUN-1999; 99US-0140354.
PR   24-JUN-1999; 99US-0140695.
PR   28-JUN-1999; 99US-0140823.
PR   29-JUN-1999; 99US-0140991.
PR   30-JUN-1999; 99US-0141287.
PR   01-JUL-1999; 99US-0141842.
PR   01-JUL-1999; 99US-0142154.
PR   02-JUL-1999; 99US-0142055.
PR   06-JUL-1999; 99US-0142390.
PR   08-JUL-1999; 99US-0142803.
PR   09-JUL-1999; 99US-0142920.
PR   12-JUL-1999; 99US-0142977.
PR   13-JUL-1999; 99US-0143542.
PR   14-JUL-1999; 99US-0143624.
PR   15-JUL-1999; 99US-0144005.
PR   16-JUL-1999; 99US-0144085.
PR   16-JUL-1999; 99US-0144086.
PR   19-JUL-1999; 99US-0144325.
PR   19-JUL-1999; 99US-0144331.
PR   19-JUL-1999; 99US-0144332.
PR   19-JUL-1999; 99US-0144333.
PR   19-JUL-1999; 99US-0144334.
PR   19-JUL-1999; 99US-0144335.
PR   20-JUL-1999; 99US-0144352.
PR   20-JUL-1999; 99US-0144632.
PR   20-JUL-1999; 99US-0144684.
PR   21-JUL-1999; 99US-0144814.
PR   21-JUL-1999; 99US-0145086.
PR   21-JUL-1999; 99US-0145088.
PR   22-JUL-1999; 99US-0145085.
PR   22-JUL-1999; 99US-0145087.
PR   22-JUL-1999; 99US-0145089.
```

```
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146385.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
alignment_scores:
Quality: 1457.00      Length: 466
Ratio: 3.865          Gaps: 4
Percent Similarity: 80.901 Percent Identity: 62.232
alignment_block:
US-09-759-990-1 x AAG22956 ..
Align seg 1/1 to: AAG22956 from: 1 to: 489
19 ACTGGTCTCCATTTCGAGTACAGAAATTCGCCGACATCAACCTCCATCTTCT 58
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
30 ThrSerSerGlyArgGluTyrLysValLysAspMetSerGlnAlaAspPh 46
69 CGCGCGTAAGGAACCTTACCCTTCTGAGGAAGAAATGCCAGGTCTTATGG 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 eGlyArgLeuGluLeuGluAlaGluValGluMetProGlyLeuMeta 63
119 TTCTTCGTGAGCGTTATTCGCTTCTTAAGCCATTCGAAGGTGTCAATC 168
::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 laCysArgThrGluPheGlyProSerGlnProPheLysGlyAlaArgile 79
169 TCTGGTTCCTCCACATGACAGTCCAGACAGCGCTCCTCATCGACACT 218
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 ThrGlySerLeuHisMetThrIleGlnThrAlaValLeuIleGluThrLe 96
219 CACAGCTCTTGGTCTGATGTGATGGCTTCTCCTGCAACATCTTCTCTA 268
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 uThrAlaLeuGlyAlaGluValArgTyrCysSerCysAsnIlePheSerT 113
269 CACAAGATACAGCGCTGCTGCTATTCGTTGCGGCCCAACAGGCACACCA 318
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 hrGlnAspHisAlaAlaAlaAlaAlaAlaArgAspSerAlaAla..... 127
319 GAGNAGCCAGCGGTATCCCACTCTCCGCTGCTGGAAGGGGGAACACTCCC 368
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
128 .....ValPheAlaIrrLysGlyGluThrLeuG1 137
369 AGAATACCTGGGAGAACATACCGCGCTCTCACATGGCCAGATGGTCAAG 418
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 nGluTyrTrpTrpCysThrGluArgAlaLeuAspTrpGlyProGlyGlyG 154
419 GCCACAGCAGGTGTGCGATGATGGTGTGATGCTACACTCTCATCTCC 468
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 lyProAspLeuIleValAspAspGlyGlyAspAlaThrLeuLeuIleHis 170
469 AAGGCTTCGAA.....TTCGAACAGCGCGTCTGCTGTTCCAGA 506
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
171 GluGlyValLysAlaGluGluIlePheGluLysThrGlyGluValProAs 187
507 GCCAACAGAGCTGACAACTCGAATACCGCTGCGTCTTCTGCTACTCA 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 pProThrSerThrAspAsnProGluPheGlnIleValLeuSerIleIleL 204
557 AGCAGGTCTTCAACCAAGACAGAACCACTGGCACACACTTGTGCGCGG 606
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 ysGluGlyLeuGlnValAspProLysLysTyrHisLysMetLysGluArg 220
```

607 ATGAACGGTGTTCCTCCGAGAGACACAACAGGTGCCACGCCTTACC 656
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
221 LeuValGlyValSerGluGluThrThrThrGlyValLysArgLeuTyrgl 237

657 GCTCGAAGAGGGGCAAACTCCTCTCCAGCCATCAACGTCACGACG 706
nMetGlnGlnAsnGlyThrLeuLeuPheProAlaIleAsnValAsnAsps 254
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

707 CTGTGTTAAAGTCCCAAGTTCGATAACACTCTACGGCTCGCGCCACTCCCTT 756
::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
254 erValThrLysSerLysPheAspAsnLeuTyrgLyCysargHisSerLeu 270

757 ATCGCATGGTATCAACCGTGCTTCGATGTCATGATCGCGGCGCAAGACG 806
||||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
271 ProAspGlyLeuMetArgAlaThrAspValMetIleLaGlyLysValal 287

807 TCTCGTCATGGGTTACGGCGATCGCGCAAGGGTCGCTCAATCCCTCC 856
eValIleCysGlyTyrgLyAspValGlyGlyGlyCysAlaLaIamMetL 304
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

857 GTGGCCAAGCGCTCGCGTTATCATCACAGAAGTCCAGCCAACTCTCGCT 906
:: ystHraLaGlyAlaArgValIleValThrGluIleAspProileCysAla 320

907 CTCACGGCTGCCATGAAGGTCACAGGTCGCGCGCATCGAGGAAGTCGT 956
LeuGlnAlaLeuMetGluGlyLeuGlnValLeuThrLeuGluAspValVa 337

957 CAAGGATGTCGATATCTTCCTTACATGCACAGGAACATCGGNATCATCT 1006
l::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 lserGluAlaAspilepheValThrThrGlyAsnLysaspIlelem 354

1007 CTGTTGACATGATGCGCCACATGAAGATGAAGGTATTGTTCGGTAACATC 1056
etValAspHisMetArgLysMetLysAsnAsnAlaIleValCysasnIle 370

1057 GCCCATTCGATAACGAATGTATACAGATGCGCTCATGAATACCCAGG 1106
GlyHisPheAspAsnGluIleAspMetLeuGlyLeuGluThrTyrgProgl 387

1107 CATCAAGCACATCCCNAATCAAGCCAGATACGACATGTGGGAATCCCCAG 1156
yValLysArgIleThrIleLysProGlnThrAspArgTrpValPheProg 404
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

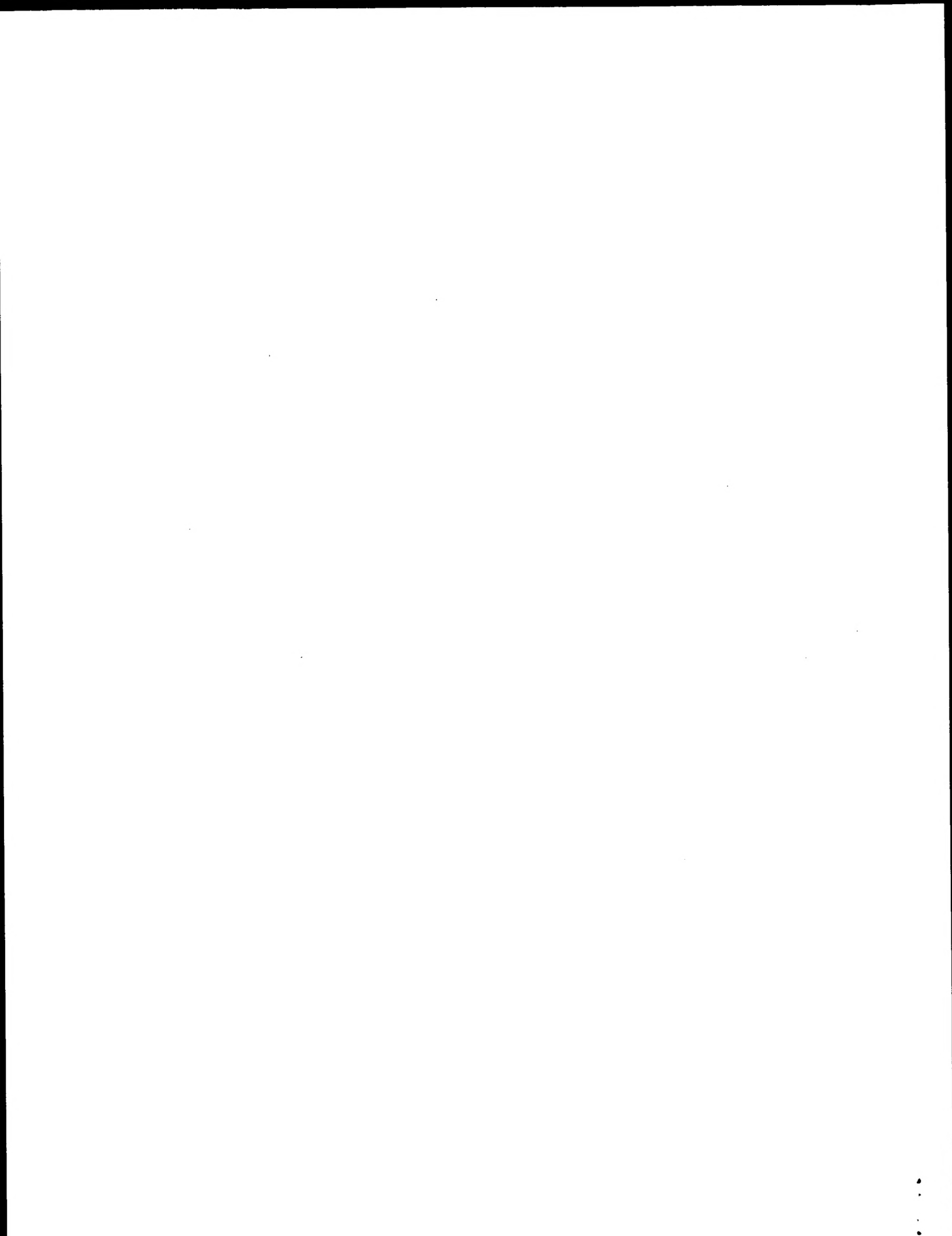
1157 ATGGCCACGCT...ATCCCTCTTCTTCTGCTCAGGCGCGCTTCTTAACTTT 1203
:: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 luThrLysalaGlyIlelleValLeuAlaGluGlyArgLeuMetAsnLeu 420

1204 GGCTCGCTACAGGTCACCCATCTTTCGTTATGTCAATGTCAATTCACAAA 1253
GlyCysAlaThrGlyHisProSerPheValMetSerCysSerPheThrAs 437

1254 CCAGACACTCGCTCAGCTCGACTCTACGAAACAGAGA.....GGAAATC 1297
lll|||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
437 nGlnValIleAlaGlnLeuGluLeutrpnasngLulysalaSerGlyLylust 454

1298 TCGAGAAGAAGGTTTACACACTTCGGAAGCATCTCGATGAAGAAGTCGCT 1347
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
454 yrGluLysLysValTyrvValLeuProLysHisLeuAspGluLysValala 470

1348 CGCCTCCACCTCGGATCTCTCGATGTCACCTTCAAAGCTTACACAG 1395
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 LeuLeuHisLeuglyLysLeuGlyAlaArgLeuThrLysLeuSerLys 486



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:55:32 ; Search time 80.77 seconds
(without alignments)
4443.117 Million cell updates/sec

Title: US-09-759-990-1

Perfect score: 1461

Sequence: 1 atggcttgcacatcacctac.....ctgatgttaccgttattaa 1461

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	490	33.5	1812	2	US-08-669-536-1
2	489	33.5	1767	3	US-08-930-894-1
3	375.4	25.7	441529	4	US-09-103-840A-1
4	223.2	15.3	2211	4	US-09-318-448-26
5	122.4	8.4	289	1	US-08-204-740-8
6	122.4	8.4	289	3	US-09-081-167A-8
7	122.4	8.4	289	3	US-09-081-395-8
8	122.4	8.4	289	4	US-09-416-833-8
9	122.4	8.4	289	5	PCT-US95-02521-8
10	105	7.2	285	1	US-08-204-740-6
11	105	7.2	285	3	US-09-081-167A-6
12	105	7.2	285	3	US-09-081-395-6
13	105	7.2	285	4	PCT-US95-02521-6
14	105	7.2	285	5	US-08-998-416-249
15	96.6	6.6	584	4	US-09-328-111-850
16	56.8	3.9	636	4	US-09-197-649-7
17	41	2.8	390	4	US-09-197-649-7
18	40.2	2.8	390	4	US-09-197-649-7
19	40	2.7	2226	2	US-08-896-005-2
20	38.4	2.6	300	2	US-08-716-942-15
21	38.4	2.6	300	2	US-08-716-942-19
22	37	2.5	4403765	4	US-09-103-840A-2
23	36.8	2.5	298	2	US-08-716-942-12
24	36.6	2.5	539	3	US-08-828-741B-12
25	36.6	2.5	539	3	US-09-160-567-12
26	36.6	2.5	599	3	US-08-828-741B-7
27	36.6	2.5	599	4	US-09-160-567-7

28 36.6 2.5 1031 3 US-08-828-741B-5 Sequence 5, Appli
29 36.6 2.5 1031 4 US-09-160-567-5 Sequence 5, Appli
30 36.6 2.5 1490 3 US-08-828-741B-3 Sequence 3, Appli
31 36.6 2.5 1490 4 US-09-160-567-3 Sequence 3, Appli
32 35.8 2.5 1229 1 US-08-440-861-1 Sequence 1, Appli
33 35.8 2.5 1229 1 US-08-433-854-1 Sequence 1, Appli
34 35.8 2.5 1229 1 US-08-174-745A-1 Sequence 1, Appli
35 35.8 2.5 1229 2 US-08-433-908B-1 Sequence 1, Appli
36 35.8 2.5 1229 2 US-08-195-947-1 Sequence 1, Appli
37 35.8 2.5 1229 2 US-08-433-885-1 Sequence 1, Appli
38 35.8 2.5 1229 4 US-08-410-614-1 Sequence 1, Appli
39 35.8 2.5 2000 4 US-09-329-350-34 Sequence 34, Appli
40 35.6 2.4 1155 2 US-08-387-942C-7 Sequence 7, Appli
41 35.4 2.4 1155 2 US-08-387-942C-20 Sequence 20, Appli
42 35.4 2.4 12588 2 US-08-387-942C-1 Sequence 1, Appli
43 35.2 2.4 28804 2 US-08-592-874-1 Sequence 1, Appli
44 35.2 2.4 28804 3 US-09-096-942-2 Sequence 2, Appli
45 35.2 2.4 28804 3 US-09-096-867-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-669-536-1

; Sequence 1, Application US/08669536

; Patent No. 5910444

; GENERAL INFORMATION:

; APPLICANT: MASUTA, CHIKARA

; APPLICANT: UEHARA, KYOKO

; APPLICANT: TANAKA, HIDEO

; APPLICANT: KAWATA, SHIGERU

; TITLE OF INVENTION: ORGANISMS IN WHICH THE EXPRESSION OF

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

; STREET: PO BOX 747

; CITY: FALLS CHURCH

; STATE: VA

; COUNTRY: USA

; ZIP: 22040-0747

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30

; APPLICATION NUMBER: US/08/669,536

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: MURPHY JR, GERALD M

; REGISTRATION NUMBER: 28,977

; REFERENCE/DOCKET NUMBER: 1254-128

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1812 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; US-08-669-536-1

Query Match

Best Local Similarity 33.5%; Score 490; DB 2; Length 1812;

Matches 887; Conservative 0; Mismatches 515; Indels 48; Gaps 4;

QY 33 cgagtacagaattgcgcacatcaccttcctgcttcgcgcgtaaggaactacccttgc 92

1152	DB	CCCTGGTGTC	AAAGAGATCA	CAATTAAGCCTCA	AAACGACAGATGGGTCTTCCTCGACAC	1211
1161	QY	ccac---gctatc	ctctcttctgctgagggccgcctcttctaa	cccttggctgctgagctacagg	1217	
1212	DB	CAACAGTGGC	ATCATTTGCTTGGCTGAGGGTCGTCTCATCAACTTGGGATGTGCCACAGG	1271		
1218	QY	tcacc	ccatcttctggtatgtcaatgtcattccacaaacacagacactcgctcagctcgacct	1277		
1272	DB	ACACCC	TAGTCTGCTGCTGCTTCACTAACCAAGTCATTGCCCAACTCGAGTT	1331		
1278	QY	ct-----acg	aaagagagaaatctctgagaaagaagatttacacactccgaagcatct	1331		
1332	DB	GTGGAATGA	AAAGACAGTGGGAAGTATGAGAAGAAAGTGTATGCTTTGGCAAAACACCT	1391		
1332	QY	cgatgaagaagtcgcctcgctccac	ctcgatctctcgatgtccacacttacaagcttac	1391		
1392	DB	CGACGAGAAGGTTGCTGCATCTTCGGA	AAGCTCGGACCGCAAGCTTACCAAACTTTC	1451		
1392	QY	acagaagcaggctgactacatcaacaggttc	caggttaggggttccttacaagcttgatgotta	1451		
1452	DB	GAAGGATCAAGCTGACTACATTAGCGTTCCAGTTGAGGGTCTCTTACAAGCCTGCTCACTA	1511			
1452	QY	ccggttattaa	1461			
1512	DB	CAGGTACTGA	1521			

RESULT 2

```

US-08-930-894-1
; Sequence 1, Application US/08930894
; Patent No. 6037524
;
; GENERAL INFORMATION:
; APPLICANT: GREENLAND, Andrew James
; APPLICANT: DRAPER, John
; APPLICANT: SKIPSEY, Marc
; APPLICANT: WARNER, Simon
; TITLE OF INVENTION: S-ADENOSYL-L-HOMOCYSTEIN HYDROLASE PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,894
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00882
; FILING DATE: 10-APR-1996
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9507381.3
; FILING DATE: 10-APR-1995
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: SHH GENE FROM ASPARAGUS
; FEATURE:
; NAME/KEY: CDS

```

Db	CGAGTACAAGGTC	AGGACATGCTC	CAGGCCGATTT	CGGCCGGCTG	GAATCGAGCTGGC	158
Qy	tgagaaggaaatg	ccaggtcttatg	ttcttcgtgagc	gttattccgcgtt	cttaagccatt	152
Db	CGAAGTGAATG	CCCTGGTCTC	ATGCGTTGTC	GACTGAA	TGTGGCCCTTCACAGCAAT	218
Qy	gaagggtgcaga	aatcttggttcc	ctccacatgac	cagtcacagtc	ccagcgcgtctctcatga	212
Db	TAAAGTGCTA	AGATTACTT	GGATCTTT	TAGATATG	ACCATTCGAATTTGATTA	278
Qy	gacatcaagctc	ttggctgcatc	atcagatggc	ttctctgcaacat	cttctctcacaca	272
Db	AACCCCTACT	CTTTGGGTGCT	GAAATAG	TGGTGTCT	TTCGAACATCTTCTCCACTCA	338
Qy	agatacagccgc	tgtgctatcg	ttgtcggccca	caggccacac	ccagagagaagccagccgg	332
Db	AGATACGGC	CGCTGCTGCC	ATTG-----	-----	-----CACGTGACAG	371
Qy	tatcccaagtct	cgctggaaggg	cgaaaacat	cccagaata	ctgggagaacacataccg	392
Db	CGCGCGCGT	GTTCCGCTG	GAAGGTG	AGACTCTG	CAGGAGTATTTGGTGTACTGAGAG	431
Qy	cgctctcacat	ggccagatg	gtcaaggcc	acacagc	aggttgtcgatgatggtgatgc	452
Db	GGCATTGACT	TGGGGTCC	AGTGTGG	CGCCGACT	TGATCGTCGACGATGGTGGTATGC	491
Qy	tacactctcatc	ctccaaaggg	cttcgaattc	gaaacagc	ggctgc-----gt	500
Db	TACACTCTG	ATTCATGAG	GGTGTTA	AGGCAAG	AAGAAAGATTTCCTAAGAAATGGGACAAT	551
Qy	tccagagccaa	gaaagctg	acaacctc	gaaacccgt	gcgttcttctgtaacctcaagca	560
Db	CCAGATCC	TAACTAC	CGATAAT	GTGAGTTT	CAGCTTACTTACTATTAATAAGGA	611
Qy	ggttcttaacca	agagaacac	taactggc	acacagtt	gtcgcgcgcatgaacggtgttc	620
Db	AAGTTGAAG	ACTGATCT	TTAAATAT	TACCAAG	TGNAGAAAGACTCGTCGCTGTTTC	671
Qy	cgaagagaa	caacaggt	gtccaccg	ctctcac	agctcgagagaggggcaactcct	680
Db	TGAGGA	AACTACCA	TGGAGT	TAAAGGCTTT	ATCAGATGCAGGCTAATGAACTTTGCT	731
Qy	cttccagccat	caacgtca	acagcgt	gttaca	aagTccaagttcgataacatctacgg	740
Db	TTCCTCTG	CTATTAAT	GTATATG	ATCTTGTT	TACCAGACAAAGTCGACAACTTGTACGG	791
Qy	ctgcgcgcact	ccctctatc	gatgggtat	caaacg	tgtcttcgatgcatgatcgcgcaaa	800
Db	ATGCGGCC	ATCCTAC	TGCCGAT	TGGTCTCAT	GAGGCTACTGATGTATGATGTGCGGAAA	851
Qy	gacagctctc	gtcatggt	ttaacggcg	atgctcg	gcaaggtcgctcaatccctcgttgg	860
Db	GTTTGCCCT	TGTTGCTG	TATGAGAT	GTCTGGCA	AGGGTTGTGTGTCCTTGAACA	911
Qy	ccaagcgctc	ggtttatcat	cacagaact	cgacccaat	ctcgtctctccaggtcgcat	920
Db	AGCCGTG	CCCCGTG	ATGTTG	ACCGAGAT	TGACCTATCTGTGCTCTCCAGGCTACCAT	971
Qy	ggaaggtcac	ccaggtccg	catcgagaa	ctctcaag	gatgcgatctctcgttac	980
Db	GGAAGGCC	TCCAGGTC	CTTAC	TCTAGAG	ATGTGCTTTCTGATGTGATATCTTGTGCAC	1031
Qy	atgcacagga	aaactgc	atatcatc	ctctgtg	acatgatggcccagatgaagataaaggc	1040
Db	CACGACCG	GTAA	CAAGGAC	ATTATCAT	TGGTTGACCACTATGGAAGATGAAGAACAAATGC	1091
Qy	tattgtcggt	taacatc	ggccatc	gtacata	cacaaattgatcacagatggcctcatgaata	1100
Db	CATTTGTT	GTGAAC	ATATGG	TCAC	TTTGCAACGAAATTCGACATGCTTTGGTCTCGAGACCTA	1151
Qy	cccagggcat	caacacat	cccatca	agccagaa	tacgacaatgtgggaattcccatgatgg	1160

LOCATION: 26..1483
OTHER INFORMATION: /codon_start= 26
US-08-930-894-1

Query Match 33.5%; Score 489; DB 3; Length 1767;
Best Local Similarity 61.1%; Pred. No. 8.9e-141;
Matches 886; Conservative 0; Mismatches 515; Indels 48; Gaps 4;

QY 33 cgagtacagaattgccgacatacaactccatgttctcgccgtaagaacttacccttgc 92
DB 61 CGAGTACAAGGTCGAAGACATGCTTCAGGCGGACTTCGGCGCCTCGAGATCAGCTGC 120
QY 93 tgagaagaaatccaggtcttatgttctctgtgaggttattccgttcttaagccatt 152
DB 121 TGAGGTCCAGATGCCAGGCTCATGGCTGCGGTGCGGTCCGAGTTTCGGCGCCGCCAGGCATT 180
QY 153 gaagggtgtcagaactctctgttccctccacatgacagtcacagacagccgtctctatcga 212
DB 181 CAAGGGCGCAAAATCACTGGATCCCTCCACATGAGATCCAAACTGCCCTCATCGA 240
QY 213 gacactcaacagctctgt 272
DB 241 AACCTTAACCGCCTCCGGCCCGGAGGTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 273 agatacagccgtctgtctatcgttgcggcccaacagggcacacagagaagccagccgg 332
DB 301 GGACCATGCCGCGCTGCCATTGCCCGTGACTCCGC----- 336
QY 333 tatccagttcttcgttggagggcgaaactcccaactcccaacttactgtggagaaacacatccg 392
DB 337 ---CTCGGTCTTCGCTTGAAGGTGAGACCTCCAGGAGTACTTGTGTGTGTGTGTGTGTGTGT 393
QY 393 cgtctcacaatggtcagatgttcaaggccacagcaggttctcgtatgtgtgtgtgtgtgtgc 452
DB 394 TGCCCTCGACTTGGGGCCCGCGGTGGCCCTGACCTCATCTGATGACGGCGGCGACAC 453
QY 453 tacactctctatctccaaagggtt-----cgaaattcgaaacagccggtgtgt 500
DB 454 CACTCTCTTGTATCATGAGGGGTGAAGCGCGAGAGGTACGAGAAGACGGGAAGAT 513
QY 501 tccagagcaacagaagctgacaaactcgaataccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 560
DB 514 GCCCGATCGGGCTCTACCGACAAATGTGAGTTCAGATCTGTCTCAATCATCAGGA 573
QY 561 ggtcttcaacagaagaacacactggcacacagttgtgtgtgtgtgtgtgtgtgtgtgtgtgt 620
DB 574 TGGGCTCAAGGTGGACCCACCACCAAGTACAGGAAGATGAAGATAGGATGTGTGTGTGTGT 633
QY 621 cgaagagcaacaaaggtgtccacgcctctaccagctcgaagagagggcgaactcct 680
DB 634 GGAGGAGACCAACCGGGTCAAGAGGCTTACCAGATCGAGGCTAACAAATTCCTTCT 693
QY 681 ctccacagcatcaacgtcaagacgtgttacaagtcgaagttcgaatcgaatcgaatcgaatcga 740
DB 694 TTTCCCTGGATCAATGTCAATGACTCCGTCACCAAGAGCAAGTTTGACATCTGTATGG 753
QY 741 ctgcccactcccttatcgtatggtatcaacgtgtctccagatgtcgtatcgtatcgtatcgtatc 800
DB 754 ATGCCGGCACTCTTCCGGATGGTCTGATGAGGGCCACTGATGTATGATGTGTGTGTGTGT 813
QY 801 gacagctctcgtatcgt 860
DB 814 GGTGTCAATGTTCTCGGTTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 873
QY 861 ccaagggcgtcgtgtatcacaagaactcgaacccaatctgcgtctcgaaggtcgaaggtcgaat 920
DB 874 GGCTGGTCCCGTGTATGTGACGGAGATCGACCCCATCTGTGTGTGTGTGTGTGTGTGTGTGT 933
QY 921 ggaaggtcaccaggtcccgccgcatcagagaagtcgtaagaggtgtcgtatcgtatcgtatcgtatc 980
DB 934 GGAGGGTCTTCAGGTCCTCACCCCTCGAGGATGTGTCTCAGAGCCGAGATATCTTTGTATC 993

QY 981 atgcacaggaactcgatatactctctgttgacatgatggcccgatgaagggataaaggc 1040
DB 994 CACCACCGGTAAACAGGACATCATCTGTCGGACCATGAGGAGATGAAGAACAATGC 1053
QY 1041 tattcgtgtaaacatcgccacttcgataacgaataatgatagatggcctcgaataa 1100
DB 1054 CATTTGCTCAACATTTGGTTCACATTTGACAAAGAGATTGACATGCTAGGTTTGGAGACATA 1113
QY 1101 cccaggtcaagcacatcccaatcaagccagaataacagacatgtgggaattcccgatgg 1160
DB 1114 CCCTGGCATCAAGAGAATCACCATCAAGCCCACTGACCGGTGGTCTTCCCTGAAAC 1173
QY 1161 ccac---gctatccctctctctgtgagggccgctctttaaaccttgaaccttgcgtcgtacagg 1217
DB 1174 CAACACTGGTATAATTTGTTCTTGTCTGAGGCCGCACTCATGAACCTTGGGTGTGCCACTGG 1233
QY 1218 tcaccactcttctgttatgtcaatgtcaattcaacaaacacacactcgtcgtcgtcgtcgtcgt 1277
DB 1234 TCACCCAGCTTTGTCTATGCTCTGCTCTCTTCCACCAACAGGTGATTGCTCAGCTAGAGTT 1293
QY 1278 ctacgaaaag-----agaggaatctcgagaagaaggtttacacacttccgaagcatct 1331
DB 1294 GTGGAATCAGAAGGCAAGCGCAAGTATGAGAAGAAGGTTTACGTGCTCCCAAGCATCT 1353
QY 1332 cgtagaagaatcgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1391
DB 1354 TGATGAGAAGTAGCAGCGCTTCACTTGGCAAGCTCGGAGCCAAAGCTTACAAAGCTCAG 1413
QY 1392 acagaagcaggtgactacatacaagttccagttcaggttgaggttccttacaagttcgtatgctta 1451
DB 1414 CCCCTCACAGCGGCACTACATCAGGCTCCCATCGAGGGTCCCTCAACGCCACCTCACTA 1473
QY 1452 ccgttatta 1460
DB 1474 CAGGTACTA 1482

RESULT 3

US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 25.7%; Score 375.4; DB 4; Length 4411529;
Best Local Similarity 61.3%; Pred. No. 1.4e-103;
Matches 879; Conservative 0; Mismatches 551; Indels 3; Gaps 1;
QY 32 tcgagtacagaattgccgacatacaactccatgttcttcggtcgaagaaactacccttg 91
DB 3629588 TCGATTTAAGATCGCCGACCTGTCTACATGCGGATTTTCGGCCGCAAGAACTCCGGATCG 3629529
QY 92 ctgagaaggaatgccaggtcttattgttcttctgtgagcgttattccgttcttcttaagccat 151
DB 3629528 CCGAGCAGGATGCGCGCCTGATGTGCTGCGCGGAGTATGCCGAGGTGCAACCC 3629469

[illegible]

Db 3628388 TCGTGATGAGCAACAGCTTCGCTAACCCAGAGCATGCCCAGATCGAGCTGGGCCAAGA 3628329
QY 1289 gaggaatactcagaagaaggtttcacacttccgaagcatctcgatgaaagtgcgtc 1348
Db 3628328 ACGACGAGTACGACAACGAGGTGTACC GGCTGCCCAAGCACCTCGACGAGAAAGTTGGCTC 3628269
QY 1349 gctccacctcggattctctcgatgtcccacttacaaaagcttcacagaagcaggctgact 1408
Db 3628268 GAATCCATGTGAGGGCCCTTGGCGGTACCTGCACCAAGCTGACCAAGGAGCAGGCCGAAT 3628209
QY 1409 acatacaactccagtgtaggtccctacaagctgatgtaccgtttattaa 1461
Db 3628208 ACCTGGCGTCTGACGTCGAGGGTCCTACACAGCCGACCACTACCGCTACTGA 3628156

RESULT 4
US-09-318-448-26
; Sequence 26, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-26

Query Match 15.3%; Score 223.2; DB 4; Length 2211;
Best Local Similarity 55.3%; Pred. No. 1.1e-58;
Matches 553; Conservative 0; Mismatches 343; Indels 104; Gaps 2;

563	QY	tcttcaaccaagcaagaacccactggcacacagttgtgcggcgcatgaagcgtgtttccg	622
451	Db		510
451	Db	tcaccaacctcatccacacaaagtaccgcagcttgcagggtatccgaggtcatctcg	510
623	QY	aagagacaaacaacaggtgtccacgcctctaccagctcgagaagagggaacatcctct	682
511	Db		570
511	Db	aggagaccagctgggtccacaacctctacaagatgatggccaatggggtatctcaagg	570
683	QY	tccagccatacaacgtcaaacgacgtgttcaaaagtcacagttcogataacatctcacgct	742
571	Db		630
571	Db	tgcctgcatacaatgtcaatgactcctccacagagcaagttagacaacctctatggct	630
743	QY	gccgcgaactcccttatogatggttatcaacctgttccogatgtcatgcgcggccaaga	802
631	Db		690
631	Db	gccggagtcctcatagatggatcaagcgggccacagatgtgatgtgccggcaagg	690
803	QY	cagcttcgtcatgggttacgcgcgatgcggcaagggctgcctcaatccctccgtggcc	862
691	Db		750
691	Db	tagcgggtggtagcgggtatggtgatgtgggcgaaggctgtgccagccctgcggggtt	750
863	QY	aaggcgtcgcggttatcatcacagaactcgaccacaatgcgctctcagggtgcctatgg	922
751	Db		810
751	Db	tcggagcccgctcatcatcacccagagattgacctatcaacgacgtcagggtgcctatg	810
923	QY	aaggctaccaggtccgcgcacaggaagtcgtcaaggatgcgatctatctcgttacat	982
811	Db		870
811	Db	agggtatggtgtaccacatggatgggctctcaggagggggaacatctttgtcacca	870
983	QY	gcacaggaaactgcgatcatcatctctgttgacatgatg-----	1020
971	Db		930
971	Db	ccagagacttatcaactatcttgcgcgttaggtgcacatagggggtcccgggag	930

```
QY 1021 ----- 1020
Db 931 tgaggagaggagagagtggtggagagctttctgtccctgacaatcccaagctcttg 990
QY 1021 -----gcccagatgaagataagctattgtcggttaacatcgcca 1061
Db 991 gctgctgacagggcaacttttgagcagatgaagatgacatctgtgttaacattggaca 1050
QY 1062 cttegataacgaataattgatacagatggcctcatgaataaccaggacatcaagcacatccc 1121
Db 1051 ctttgacgtggagatcgatgtaagtggctcaacagagaac---gccgtggagaagtgaa 1107
QY 1122 aataaagccagaataacacagatggtgaattccacagatggcagcgtatctctctttgc 1181
Db 1108 cataaagccaggtggaccgtgttcgttggaagatggcgccgcatcatctctgtgc 1167
QY 1182 tgaggggccgcttttaaccttggctgcgtacaggtacacacatctttctgtatgcaat 1241
Db 1168 cgagggtcggtgttcaacctgggtgttgccatggccacccagcttctgtgatgataa 1227
QY 1242 gtcattcacaaacagacactcgtctcagctcagctcgtacgaaagaggaatctcga 1301
Db 1228 ctctctcaccaaccaggtgtggtgcagatcgagctgtgacccatccagacaagtaacc 1287
QY 1302 gaagaaggtttacacacttccgaagcatctcgatgaagaagtcgctcgaacatcc 1361
Db 1288 cgttgggttcattctctgcccgaagctggatgagcagtggtggaagccacactgg 1347
QY 1362 atctctgatccacacttccaaagcttacacagaagcaggctgactacatcaacgttcc 1421
Db 1348 caagctgaatgtgaagtgaacaaagctaaactgagaagcaagccagtaacctggaatgc 1407
QY 1422 agtgagggtccctacaagctgtgatgttaccgtttataa 1461
Db 1408 ctgtgagtcgcccctcaagcggatcatcactaccgctactga 1447

RESULT 5
US-08-204-740-8
; Sequence 8, Application US/08204740
; Patent No. 5753432
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alligretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,740
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5753432nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
```

```
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-204-740-8
```

```
Query Match 8.4%; Score 122.4; DB 1; Length 289;
Best Local Similarity 71.1%; Pred. No. 3.7e-28;
Matches 162; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 37 tacacaattgcgacatacaacctccatgtctctcggcgttaaggaacttacccttgctgag 96
Db 62 TACAAGTCGCCGACATCGCGCTGGCTGGGAGACGCAAGGCCCTGGACATTGCTGAG 121
QY 97 aaggaaatgccaggtcttattgttcttcgtgagcgttatccgccttcttaagccattgaag 156
Db 122 AACGAGATGCCGGGCTGATGCTATCCGGGACGGGTACTCGGGCTCCAGCCACTGAAG 181
QY 157 ggtgtcagaatctctgtgttccctccacatgacagtcagtcagacagccgtctctcatogagaca 216
Db 182 GCGGCCCGCATCGCTGGCTGCCTGCACATGACCGTGGAGACGGCGCTCCTCATTTGAGACC 241
QY 217 ctacagactcttggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 264
Db 242 CTCGTACCCCTGGGTGCTGAGGTGTCAGTGTCCAGCTGCCAATCTTTC 289
```

```
RESULT 6
US-09-081-167A-8
; Sequence 8, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,167A
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083745nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-081-167A-8

Query Match 8.4%; Score 122.4; DB 3; Length 289;
Best Local Similarity 71.1%; Pred. No. 3.7e-28;
Matches 162; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 37 tacagaattccgacatcaacctccatgttctcgccgtaaggaacttacccctgtctgag 96
DB 62 TACAAAGTCGCCGACATCGGCTGGCTGCTGGGAGCGCAAGGCCCTGGACATTGCTGAG 121

QY 97 aaggaatccaggtcttattgttcttcgtgagcgttattcccgcttcttaagccattgaag 156
DB 122 AACGAGATGCCGGCCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAGCCACTGAAG 181

QY 157 ggtgtcagaatctctgttctccctccacatgacagtcagagccgctctctcatcgagaca 216
DB 182 GGGCCCGCATCGCTGGCTGCTGCATGACCGTGGGAGCGCGCTCTCATTGAGACC 241

QY 217 ctcaagctctgtgtctgagtcagatggggtctctctgcaacatcttc 264
DB 242 CTCGTACCCCTGGGTGCTGAGGTGCAGTGGTCCAGCTGCAACATCTTC 289

RESULT 8
US-09-416-833-8
; Sequence 8, Application US/09416833
; Patent No. 6197521
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Alligretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,833
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/204,740
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6197521nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: .93,354-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-416-833-8

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-081-167A-8

Query Match 8.4%; Score 122.4; DB 3; Length 289;
Best Local Similarity 71.1%; Pred. No. 3.7e-28;
Matches 162; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 37 tacagaattccgacatcaacctccatgttctcgccgtaaggaacttacccctgtctgag 96
DB 62 TACAAAGTCGCCGACATCGGCTGGCTGCTGGGAGCGCAAGGCCCTGGACATTGCTGAG 121

QY 97 aaggaatccaggtcttattgttcttcgtgagcgttattcccgcttcttaagccattgaag 156
DB 122 AACGAGATGCCGGCCCTGATGCGTATGCGGAGCGGTACTCGGCCCTCCAAGCCACTGAAG 181

QY 157 ggtgtcagaatctctgttctccctccacatgacagtcagagccgctctctcatcgagaca 216
DB 182 GGGCCCGCATCGCTGGCTGCTGCATGACCGTGGGAGCGCGCTCTCATTGAGACC 241

QY 217 ctcaagctctgtgtctgagtcagatggggtctctctgcaacatcttc 264
DB 242 CTCGTACCCCTGGGTGCTGAGGTGCAGTGGTCCAGCTGCAACATCTTC 289

RESULT 7
US-09-081-395-8
; Sequence 8, Application US/09081395
; Patent No. 6083746
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Kazarov, Alexander
; APPLICANT: Mazo, Ilya
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,395
; FILING DATE: 18-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6083746nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-KK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-081-395-8

	Best Local Similarity	71.1%;	Pred. No. 3.7e-28;	Mismatches 66;	Indels	Gaps	0;
	Matches 162;	Conservative	0;				
Qy	37	tacgaattgccgacaatcaaacctcattgttctcggcgtaaagaaacttacaccttgctgag	96				
Dd	62	TACAAATCGCCGACATCGGCCTGGCTTGGTGAGACGAAGGCCGTGGACATTGCTGAG	121				
Qy	97	aaggaaatgccagygcttcatggttcttcgttagcgttatcttcgcttctaagccattgaag	156				
Dd	122	AACGAGATGCCGGGCCCTGATCGTATGCGGAGCGGTACTCGGCCCTCAAGGCACATGAAG	181				
Qy	157	gggttcagaatctctggttcccctccaatcacagtccagaccgcgtctctcatcgagaca	216				
Dd	182	GGGCCCGCATCGTGGCTGGCTTCACATACACCGTGGAGACGGCCGTCCTCATTGAGACC	241				
Qy	217	ctcaacgctcttggtgatgctgagatgggcttctctgcaacatcttc	264				
Dd	242	CTGCTCACCCCTGGGTGTTGAGGTGCAGTGGTCCAGTGCACATCTTC	289				

```

RESULT          9
PCT-US95-02521-8
; Sequence 8, Application PC/TUS9502521
; GENERAL INFORMATION:
;
; APPLICANT:
;
; TITLE OF INVENTION: Methods for Identifying Genetic
; TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
; TITLE OF INVENTION: Growth in Cancer Cells
; NUMBER OF SEQUENCES: 13
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02521
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
PCT-US95-02521-8

```

Query Match	8.4%;	Score 122.4;	DB 5;	Length 289;
Best Local Similarity	71.1%;	Pred. No. 3.7e-28;		
Matches 162;	Conservative 0;	Mismatches 66;	Indels 0;	Gaps 0;
37	tacagaattgcgcacatacaactccatgttctcgccgctaaggaacttacccctgtcgag	96		
Db				
62	PACAAAGTCGCGACATCGGCTGGCTGGCGGACGCAAGGCCCTGCACATTGCTGAG	121		
QY				
97	aaggaaatgccaggtctctatggtttctgtgagcgttattccgcttctaagccattgaag	156		
Db				
122	AACGAGATGCCGGGCGCTGATCGTATGGGGAGGCGGTACTCGGCCCTCCAAGCCACTGAAG	181		
QY				
157	gggttcagaatctctggttccctccacatgacgttccagacagcgcgtctccatcgagaca	216		
Db				
182	GGGCGCCGCATCGTGGCTGCCATGACCGTGGAGAGCGCGCTCCTCATTTGAGACC	241		
QY				
217	ctcacagctctggtcgtatgtcagatggcgcttccctgcacaactcttc	264		
Db				
242	CTGTCACCCCTGGGTCGTGAGGTGCAGTGGTCCAGCTGTCAACATCTTC	289		
QY				

RESULT 10
US-08-204-740-6
: Sequence 6, Application US/08204740
: Patent No. 5753432
GENERAL INFORMATION:

APPLICANT: Gudkov, Andrei
 APPLICANT: Kazarov, Alexander
 APPLICANT: Mazo, Ilya
 APPLICANT: Roninson, Igor B
 TITLE OF INVENTION: Methods for Identifying Genetic
 TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
 TITLE OF INVENTION: Growth in Cancer Cells
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Allgretti & Witcoff, Ltd.
 STREET: 10 S. Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/204,740
 FILING DATE: 04-MAR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5753432nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 93,354-C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELEX: 910-221-5317
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 285 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 US-08-204-740-6

Query Match	7.2%;	Score 105;	DB 1;	Length 285;
Best Local Similarity	64.7%;	Pred. No. 8.6e-23;		
Matches 156;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;
QY	24	tgtccattcgcagtacagaattgccacatcaacctccattcttcgcgcgtgaagaacct	83	
Db	45	TGATAAATGCCCTACAAGCTCGGACATCGGACATCGGCCGCTGGGGACGGAAGCTCT	104	
QY	84	tacccttgctgagaagaaatgccaggtcttatggttcttcgtgagcgtttaticccgcttc	143	
Db	105	GGATATAGCTGAGATGAGATGCAGGGTGTGATGCGCATCGGGAGATGTACTACGCCCTC	164	
QY	144	taagccattgaagggtgtcagaaatctctggttccctccacatgacagtcacagacagcgt	203	
Db	165	CAAGCCACTGAAGGGTGCTCGGATTGCTGGCTCCCTGCGCATGCCGTGGAGACTGCTGT	224	
QY	204	cctcatcgagacactcacagctcttggtgctgatgtcagatgggtctcctgcaacatctt	263	
Db	225	TCTCATTTGAGACTCTCGTGCCCTGGGTGCTGAGGCGCGTGGTCCAGTGC AACATCTT	284	
QY	264	c 264		
Db	285	c 285		

RESULT 11
US-09-081-167A-6
; Sequence 6, Application US/09081167A
; Patent No. 6083745
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei

APPLICANT: Kazarov, Alexander
 APPLICANT: Mazo, Ilya
 APPLICANT: Roninson, Igor B
 TITLE OF INVENTION: Methods for Identifying Genetic
 TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
 TITLE OF INVENTION: Growth in Cancer Cells
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 STREET: 300 S. Wacker Drive, 32nd Floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/081,167A
 FILING DATE: 18-MAY-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6083745man, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 93,354-KK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-913-0001
 TELEFAX: 312-913-0002
 TELEX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 285 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-081-167A-6

	Query Match	7.2%	Score 105;	DB 3;	Length 285;
	Best Local Similarity	64.7%;	Pred. No. 8.6e-23;		
	Matches 156;	Conservative	0;	Mismatches 85;	Indels 0; Gaps 0;
QY	24	tgtctcattgcagtagacagaattgccgacatcaacctccatgtcttcgcgccgtaaggaact	83		
DB	45	TGATAAATCCCTCCCTAAAAGTCGGGACATCGGACTGGCCCTGGGGACGGAGGCTCT	104		
QY	84	tacccttgctgagaaggaatgccagggtcttatggtttcttgtagcgttattccgcctc	143		
DB	105	GGATAATAGCTGAGATGAGATGCCAGGGTGTGATGCGCATGCGGGAGATGACTCAGCCTC	164		
QY	144	taagccattgaagggtgtcagaattcttggttccctccatcagatcacagacagaccgt	203		
DB	165	CAAGCCACTGAAGGGTGCTCGCATTCCTGGGTCCTGCGCATGACCGTGGAGACTGCTGT	224		
QY	204	cctcatcagacactcacagctcttggtgtgtagtgcagatggggttcttcgcaacatctt	263		
DB	225	TCATATTGAGACTCTCGTGGCCCTGGGTGCTCAGGCGCGGTGGTCCAGTGCACATCTT	284		
QY	264	c 264			
DB	285	c 285			

RESULT 12
 US-09-081-395-6
 ; Sequence 6, Application US/09081395
 ; Patent No. 6083746
 ; GENERAL INFORMATION:
 ; APPLICANT: Kazarov, Andrei
 ; APPLICANT: Kazarov, Alexander

us-09-759-990-1.rni

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/998,416
  FILING DATE: 24-DEC-1997
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: CH 0016/97
    FILING DATE: 31-DEC-1996
  ATTORNEY/AGENT INFORMATION:
    NAME: Meigs, J. Timothy
    REGISTRATION NUMBER: 38,241
    REFERENCE/DOCKET NUMBER: PF/5-30306/A/GCGC1976
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 919-541-8587
    TELEFAX: 919-541-8689
  INFORMATION FOR SEQ ID NO: 249:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 584 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    ORIGINAL SOURCE:
      ORGANISM: PAG1221UP
      US-08-998-416-249

```

	Query Match	6.6%	Score 96.6;	DB 4;	Length 584;	
	Best Local Similarity	53.4%;	Pred. No. 5.3e-20;			
	Matches 236;	Conservative	0;	Mismatches 203;	Indels 3;	Gaps 2;
Qy	890	tcacccaactcgcctccaggtccagtccatggaaggtaccagggtccgcgcgatcgagg	949			
Dd	3	TCGACCTATCAAGCCCTTGCAGCGCTGCTATGGAGGGCTATCAGTCACCACCTATGGACC	62			
Qy	950	aagtcgtcaaggatgctcgatatcttcdttacaatgcacaggaataatcgatacatctctg	1009			
Dd	63	AGTGGCCAGCTACGGGCCAGGTTTTGTGTCAACCACCGCGGTGCAGANACATCATCAAGA	122			
Qy	1010	ttagcatgatggcccagaataaagataaaggctattgtcgttaaatcagggcaccttcgata	1069			
Dd	123	AGGAGCACTCTTTGGCCATTGCTTGAGNGCCATTGTGTGAACAATCGGCCACTCGACA	182			
Qy	1070	acgaattgatacacagatggcctcatgaaataccccaggatcaaacacatcccaatcaagc	1129			
Dd	183	TCGAGATCGAGCTGCGCTTGGCTAAAGGCCAACGGCGTCCGANGCGGTCAACATTAAAGCCAC	242			
Qy	1130	cagaatacacgacatgtggaatccacagatgccagctatctcctcttctgtcagggcc	1189			
Dd	243	AAGTCGACCGGTACTTGTTTCCTCCGCGCACACAGGTCACTCCTGCTGCCGAT--GGTTA	300			
Qy	1190	gccttctttaaccttggtgcgtacaggtcac-coatcttttgttatgtcaatgtcaatc	1248			
Dd	301	GACTAGTCAACCTTAAGCTGTGCCACTGGGCCACTCTCCGTTTGTGCATGCTGTGCTTTTC	360			
Qy	1249	acaaacagacacatcgctcagctcgaccttacgaaaagagagaaaaatctcgagagaag	1308			
Dd	361	TCCAACCAAGTCTTTGGCACAGATGTCNTTCAAGGGCAATNAAGAGCCCTTCAAANA	420			
Qy	1309	gtttacacacttcggaagcatc	1330			
Dd	421	ATTNNNTNNTTCCCAAAAAAC	442			

Search completed: August 17, 2002, 05:49:35
Job time: 14043 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:55:57 ; Search time 1886.52 Seconds
(without alignments)
10452.607 Million cell updates/sec

Title: US-09-759-990-1
Perfect score: 1461
Sequence: 1 atggttgcaaatcaactac.....ctgatgttaccgttattaa 1461

Scoring table: IDENTITY_NUC
Gapex 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estio: *
8: em_hic: *
9: gb_esti: *
10: gb_est2: *
11: gb_hic: *
12: gb_gss: *
13: em_gss_hum: *
14: em_gss_inv: *
15: em_gss_pln: *
16: em_gss_vrc: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442.2	30.3	1550	10	BM321022
2	343	23.5	1076	10	BG37961
3	339.4	23.2	776	10	BM321423
4	338.8	23.2	843	10	BF256594
5	338.2	23.1	951	10	BM321451
6	333.6	22.8	800	10	BM321451
7	330.2	22.6	977	9	AL163774
8	329.2	22.5	1163	10	BI244550
9	316.2	21.6	908	10	BF263419
10	313.6	21.5	867	10	BF253807
11	311	21.3	809	10	BE559309
12	307.4	21.0	804	10	BM411809
13	303.4	20.8	916	10	BG444868
14	302.8	20.7	769	10	BF863943
15	301.2	20.6	844	10	BF256768
16	300	20.5	730	10	BM112072
17	299.8	20.5	802	12	BH604363

c	18	299.6	20.5	732	10	BI750640
	19	299.2	20.5	862	10	BM321023
	20	299	20.5	726	10	BE413441
	21	298.8	20.5	689	9	AL503944
	22	297.2	20.3	932	10	BG366184
c	23	296	20.3	902	12	AZ679560
	24	295	20.2	1004	9	AL518398
	25	292.6	20.0	628	10	BI874978
	26	292.2	20.0	738	10	BI405371
	27	291.8	20.0	1070	9	BE033415
	28	291.4	19.9	906	10	BF258750
	29	290	19.8	676	9	AV916518
	30	290	19.8	774	10	BG599587
	31	289.2	19.8	904	10	BG444506
	32	288.8	19.8	672	10	BG599809
c	33	288.4	19.7	938	12	BH160428
	34	287.6	19.7	1050	9	BE036599
	35	286.4	19.6	893	10	BG478112
	36	286	19.6	908	10	BM453763
	37	285.8	19.6	648	10	BG274691
c	38	285.6	19.5	707	10	BI751148
	39	284.8	19.5	657	10	BI959820
	40	283.8	19.4	607	10	BF253757
	41	283	19.4	741	10	BM412612
	42	282.4	19.3	957	10	BF268594
	43	282	19.3	821	10	BF261131
	44	281	19.2	625	10	BM323515
	45	280.2	19.2	828	10	BI259193

ALIGNMENTS

RESULT 1

LOCUS

DEFINITION

BM321022 1550 bp mRNA linear EST 03-JAN-2002
rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC 3.3.1.1), mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 721; Conservative

Score 442.2; DB 10; Length 1550;

Pred. No. 6.3e-106;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

Mismatches 407; Indels 10; Gaps 3;

QY 324 gccagccggtatccagctcttcgcttgcgaggaagcgaacactccccaataacttgagagaa 383
Db 44 GCAGCCGCGCTCTCGGCTTCGCTTGAAGGGCGGAGAACCTCCAGAGTACTGGAGTG 103
QY 384 cacataccgctctcacatggccagatggttcaaggccacagcaggttgcgatgag 443
Db 104 CACCTGGAGGCGCTTCTGCTGGCGCCCTACCAAGGGCCCTCAGATCATCTCGACGACGG 163
QY 444 tggtagtactacactcttcacatcccaagggttcgaattcgaacagcgggtgctgtcc 503
Db 164 CGGTGACGCGACTCTGTATGATCACAAGGGGTTCGGCGGCGAGGACAACCCCAAGCTGCT 223
QY 504 agagcaacaagaagctgacaactcctgaataccgctgtcttctgtacactcaagcaggt 563
Db 224 CGAG---GACGACGAGGGCTCGAGGAGTGCCTTCCCTCAACAACGCTGCTCAAGCAGGT 280
QY 564 cttaaccaagaagaaccactggaacacagttgttgcgggcatgaacggtgtttccga 623
Db 281 CCAGAAGGAGGAGCGCGGCTTCTGGCACAAGATCTCTCCCGAGATCGCGGTGTGACGGA 340
QY 624 agagacaacaaggtgtccaccgctctacagctcgagagagggcgaactccctctt 683
Db 341 GGAGACGAGCTGCGGTGATGAGGTGTACAGCTGTCCCGGACGCGACGCTGCTGCTGTT 400
QY 684 ccagccatcaacgtcaacgagcgtgttacaagtccaagtcttcgataacatctacggtg 743
Db 401 CCGGCGCTCAACGCTCAAGGACTCTGTCCACCAAGAGCAAGTTTGACAACTCTACGGCTG 460
QY 744 ccgcaactcccttatcgatggttatcaacacgctgtcttcgatacgtatgcggcgaagac 803
Db 461 CCGCCACTCGCTCATCGAGCGGATCAAGCGCGGCGACCGAGTGTGCTGCGGCGCAAGGT 520
QY 804 agctctcatgattacgagcgtgtcgcaaggtcgctgcctcaatccctccgtggca 863
Db 521 CCGCTGTGCGGGGTACGGGACGCTGGCAAGGGCTGGCCGCGAGTGTGCTGCGCGGCCA 580
QY 864 aggcctgcgcttatcatcacagaactcgacccaatctgcgtctctccagcgtgpcatgga 923
Db 581 GGGCTGCGCGCTCATGTCAGCGAGATCGACCCCATCTGCGCGCTGCGAGGCTGATGTCG 640
QY 924 aggtacacaggttcgcgcatgaggaagtcgtcaagagatgcgatacttcgtacatg 983
Db 641 GGGCTGCGAGGTCAACACGCTCGAGCGGGGCTGCGCCGCGGAGATCTTCGTGACGGC 700
QY 984 cacaggaactcgatcatctctgtgacatgtagccagatgaggaataaggctat 1043
Db 701 GACGGGCAACCAACATCATATGCGCGAGCATGAGCAAGATGCGCCACAACCTCGGT 760
QY 1044 tgcggttaacatcgccacttcgataaacgaattgtatcacagatggcctcatgaataccc 1103
Db 761 CGTCTGCAACATCGGCGCACTTTGACAAACGAGATCGAGTGTGCGGGGCTCAGGAAGTGGCC 820
QY 1104 aggcataagaacatcccaatcaagcagatacagcatgtgggaatcccgatggcca 1163
Db 821 CGGCGTCAAGTGGCTCAACATCAAGCGCGAGTGGAGTACGATCTTTGCGGCGGCCA 880
QY 1164 cgtatctctcttctgtagggcgctcttcttaaccttggctgcgatacaggtcaccc 1223
Db 881 CGGCATCATCTGCTGCGGAGGGCGGCTGGTCAACCTCG-TGCGGCAAGGGCCACCC 939
QY 1224 atcttctgtagtaagtgcattcaacaacacagacactcgtcagctcgaccttaaga 1283
Db 940 CAGNTTGTGATGNNNACAGCTTCAACAACACAGAGCTGGCCCGAGNTTAAAGTCTG--- 996
QY 1284 aaagagagaatctcgaagaaggtttacacacttccgaagcatctcgatgaagaagt 1343
Db 997 ---GCCGAGAAAGTGGACATCGCCGTCATCACCATCCCAAGATCTTCGAGGAGGT 1053
QY 1344 cgtcgctccacactcgatctctcgatgtccaccttacaaagtttacacagaagcagggc 1403
Db 1054 CGCCCGCTTGCACCTCGAAGAGCTCAAGGTGCTGNTCACCAGCTGACCCGCGCCAGGC 1113

QY 1404 tgactacatcaacgttccagttgaggtccttacaagctgtgagcttaccggttataa 1461
Db 1114 CAAGTACATGACGCTCCCGTGCAGCGCGCTTACAAGCGCGCCACACTACCGCTANTAA 1171

RESULT 2
BG837961 1076 bp mRNA linear EST 25-MAY-2001
LOCUS Zm10_03e04_A Zm10_AAFPC_ECORC_Fusarium_graminearum_corn_silk Zea
DEFINITION mays cDNA clone Zm10_03e04, mRNA sequence.
ACCESSION BG837961
VERSION BG837961.1 GI:14204284
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1076)
AUTHORS Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott
D. and Tinker,N.A.
TITLE Expressed Sequence Tags from Maize Silk Six Hours After Silk
Channel Inoculation with Fusarium graminearum
JOURNAL Unpublished (2001)
COMMENT Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: harrisli@em.agr.ca.
FEATURES
source
1..1076
Location/Qualifiers
/organism="Zea mays"
/cultivar="CO388"
/db_xref="taxon:4577"
/clone="Zm10_03e04"
/clone_lib="Zm10_AAFPC_ECORC_Fusarium_graminearum_corn_silk"
/tissue_type="Silk"
/dev_stage="4-5 days post-silk emergence"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
Site 2: XhoI; Field-grown corn was silk channel-inoculated
in the morning (~10 am) with 1 ml of a macroconidial
suspension (500,000 spores/ml) of Fusarium graminearum and
silk channels were collected and immediately frozen in
liquid nitrogen 6 hours later. RNA was extracted from
silk tissue between 1 cm below and above the inoculation
point in the silk channel. RNA from five silk channels was
pooled."

BASE COUNT 228 a 281 c 288 g 278 t 1 others
ORIGIN

Query Match 23.5%; Score 343; DB 10; Length 1076;
Best Local Similarity 67.2%; Pred. No. 9.2e-80;
Matches 533; Conservative 0; Mismatches 250; Indels 10; Gaps 3;

QY 678 cctcttcccgacatcaacgtcgaacgctgttacaagttccaagttcgataacatcta 737
Db 1 CCGTTCCTCCCTGCAATTAACGTACGATTCGTCACCAAGAGCAAGTTTGACAACCTGTA 60
QY 738 cggctgcgcgaactcccttatacgtatggtatcaacacggtgcttcggtatgctatgcggcg 797
Db 61 TGGTTGCGGCGCACTCTCCCTGATGTTGATGAGGCGCCACCGAGCTTATGATCGCCGG 120
QY 798 caagacagctctcgtcatggttacgcgagtgctgcgaagggctgcgaatccctccg 857
Db 121 TAAGTTGCGGTGCTGCGGATACGCTGATGTTGGCAAGGTTGCGCGCTGCTCACTCAA 180
QY 858 tg-gccaaggcgtcggtttatcatcacagaactcgacccaatctcgtcgtctccaggtg 916

1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 951 Std Error: 0.00
POLYA-No.

FEATURES

source Location/Qualifiers
1. 951

/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"

BASE COUNT 186 a 321 c 303 g 139 t 2 others

ORIGIN

Query Match 23.18; Score 338.2; DB 10; Length 951;
Best Local Similarity 62.28; Pred. No. 1.6e-78;
Matches 598; Conservative 0; Mismatches 330; Indels 33; Gaps 3;
QY 33 cgaatcagaattccacatcaactccatgtttctcgccgttaaggaacttacccttgc 92
DB 24 CGACTAAGGTCAAGGACATCAACTCGCGACTTCGCGCCGCAAGAGATCGAGCTGC 83
QY 93 tgagaagaaatccaggtcttattgtttcttcgtgaggttattcgcgttcttaagccatt 152
DB 84 CGAGTACGAGATCCCGGTCTGTATGGCCACCAG---GACCAAGTACGGCCCGAGTGCCT 140
QY 153 gaaggtgtcagaattctgttctccctccacatgacagtcacagcagcgtctctatcga 212
DB 141 CCAGGGCGGAGGATACCGGCTCGCTCCACATGACATCCAGACCGCGCTCTCATCGA 200
QY 213 gacactcacagcttctgtctgtatgacagtggtcttccttcgcaacatcttctctaca 272
DB 201 GACGCTCAAGTTCCTCGGCGCAAGATCCGCTGGTGTCTCGTGAACATCTTCTCGACGA 260
QY 273 agatacagccgtctgtctatcgttctgcccacagggccacacagagaagccagccgg 332
DB 261 GGACACGCGCGCGCCCATCG-----CCAGGCGCGG 293
QY 333 tatccagctctcgtctggaagggcgaaacactcccagaatactggggagacacataccg 392
DB 294 CGTCTCGGTCTTCGCTTGAAGGGCGAAGACCTCCAGGAGTACTGGGAGTGCACCTGGAA 353
QY 393 cgtctcacatggccagatggtcaaggccacacagcaggttctcgtatggtggtgatgc 452
DB 354 GGCCCTGTCTTCCGCGCCCTACACAGGCGCTCAGATCATCTCTCAGCAGCGCGGTGACGC 413
QY 453 tacactcctcatctcaagggtctgaattcgaacacagccggtgtgttccagagccaac 512
DB 414 GACTCTGATGATCCACAAGGGGTTCGCGCGCGAGGACAAACCCCAAGCTGTCTGGAG---GA 470
QY 513 agaagctgacaacctcgaataaccgctgcgttcttctgctacactcaagcaggtcttcaaca 572
DB 471 CCAGGAGGCGCTCGAGAGGTGCGCTTGCCTTACACACGTGCTCAAGCAGGTCCTCAAGAGA 530
QY 573 agacaagaaccactggcacacagttgtcgcggcatgaacggtgtttccggaagagacaac 632
DB 531 GCAGCGCGGCTTCTGGCACAAGATCTCCCGAGATCCGCGGTCTCAGCAGGAGAGACAC 590
QY 633 aacaggttccaccctctaccagctcagaagagggaaggaactctcttctccagccat 692
DB 591 GACTGGCGTATGAGGCTGTACAGCTGACCGCGAGCGGAAGCTGTGTTCGCGCGGT 650
QY 693 caacgtcaacgagcgtgttacaagtcgaagttcgtataacatctacgctgcgcgcactc 752
DB 651 CAACGTCAAGGACTCTNTCACCAGAGGAGTTCGACACATCTACCGGCTGCCGCCACTC 710
QY 753 ccttatcagtgatataaacctgtctcccgatgtcatatgtagcgggggaagacagctctcgt 812
DB 711 GCTATCTACGCGCATCAAGCGGGGACCGACGCTGATGCTCGCGGCAAGGTTCGCGCTCGT 770
QY 813 catgggttacggcgtatcgcaagggttcgctcgaatcctcgtcgtgccaagcgctcg 872
DB 872

Db 771 CGCGGCTACGGCGACGTGGGCAAGGGCTGCGCGAGTCTGCGCGCCAGGGTGC 830
QY 873 cgttatcacaagaactcgaacccaatctcgtctccagctgctccatgggaaggttacca 932
DB 831 CGTATCGTGAACGAGATCGACCCATCTGCGCGCTGCAGCGCTCGATGGCGGCTTCA 890
QY 933 ggtccgcccgcagcaggaagtcgcaagagtcgatatcttcgttatcatcacagga 992
DB 891 GGTCAACGCTCGAGCGGGGCTCGACCGCGCGAGATCTTGGTGACGGGACGGCAA 950
QY 993 c 993
DB 951 c 951

RESULT 6
BI929547
LOCUS
DEFINITION
EST549436 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
clone cTOB29E10 5' end, mRNA sequence.
ACCESSION
BI929547
VERSION
BI929547.1 GI:16243679
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
1 (bases 1 to 800)
AUTHORS
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
Unpublished (2001)
CONTACT
CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1. 800
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB29E10"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research. Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

FEATURES

source

Location/Qualifiers
1. 800
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB29E10"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research. Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

BASE COUNT
ORIGIN

Query Match 22.88; Score 333.6; DB 10; Length 800;
Best Local Similarity 65.68; Pred. No. 2.4e-77;
Matches 521; Conservative 0; Mismatches 264; Indels 9; Gaps 2;
QY 677 tctcttcccagccatcaagcgaacgctgtttacaaagtcgaagttcgaatacatc 736
DB 5 TGTCTTCCCTGATTAATGTTAATGACTCTGTACCAAGAGCAAGTTTGACACTTGT 64
QY 737 acgggtcgccgactcccttcatcgtatcgaacgctgctccgctcgtatcgtcgcg 796
DB 65 ACGGATGCCGCCACTCACTTCCCGATGGTCTCATGAGGGCTACTGATGATTGCTG 124

LOCUS B1244550 1163 bp mRNA linear EST 12-JUL-2001
 DEFINITION MIN41 MIN Nitrogen-replete Schizopyllum library Schizopyllum
 commune cDNA 5' similar to adenosyl homocysteine, mRNA sequence.
 B1244550
 ACCESSION B1244550.1 GI:14713532
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Schizopyllum commune.
 Schizopyllum commune.
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 Agaricales; Schizophyllaceae; Schizopyllum.
 1 (bases 1 to 1163)
 REFERENCE Jackson, E.N., Honaas, L., Powell, I., Hittinger, C.T., Green, A., Lilly
 W.W. and Gathman, A.C.
 TITLE Expressed sequence tags from Schizopyllum commune nitrogen-replete
 and nitrogen-limited libraries, Summer 2001
 JOURNAL Unpublished (2001)
 COMMENT Contact: Gathman AC
 Biology Department
 Southeast MO State University
 1 University Plaza, Cape Girardeau, MO 63701, USA
 Tel: 5736512361
 Fax: 5739866433
 Email: agathman@biology.semo.edu
 Seq primer: T3
 POLYA=Yes.

FEATURES
 Source Location/Qualifiers
 1..1163
 /organism="Schizopyllum commune"
 /strain="4-40"
 /db_xref="taxon:5334"
 /clone_lib="MIN Nitrogen-replete Schizopyllum library"
 /tissue_type="mycelium"
 /note="Vector: lambda Zap; Site_1: EcoRI; Site_2: XhoI;
 4-day-old mycelia of Schizopyllum commune were
 transferred from minimal (nitrogen-replete) medium to
 fresh minimal medium. RNA was extracted twelve hours after
 transfer and cDNAs prepared."
 BASE COUNT 244 a 366 c 313 g 205 t 35 others
 ORIGIN

Query Match 22.5%; Score 329.2; DB 10; Length 1163;
 Best Local Similarity 62.6%; Pred. No. 4.3e-76;
 Matches 538; Conservative 0; Mismatches 313; Indels 9; Gaps 3;
 QY 604 ggcatgaacggtttccgaagacacacacaggtgtccacgctctaccagctcgag 663
 Db 6 GACATCGCGGTATCTCTGAGGAGACCACTACCGGTGTTCACCACTGTACAAGCGCTTC 65
 QY 664 aaggaggcaactctctccagccatcaacgtcaacgacgctgttaca--aagtcca 721
 Db 66 CGGATGGCAAGTCAAGTCCCGCCATCAACGTTAACGACTCGGTACACNNGAAGTCCA 125
 QY 722 agtctgaatacatctac-ggctgcgcgcactcccttatcgatgtgtatacaacggtcttcc 780
 Db 126 AGTTCGACAACTGTACNGGCTGCCGGAGTCCCTCGTGTGATGTCATCAAGCGTCCACC 185
 QY 781 gatgtcatatcgcggaagacagctctctgctatggttaccgcatgctcggaagggc 840
 Db 186 GATGTGATGTTGCGGCAAGTTCGCGGTGTCTGCTGGGTTCGGTGTGATGTCGCAAGGCG 245
 QY 841 tgcgtcaatccctcgtggccaaagcgctcggtttatcatcacagaactgacccaatc 900
 Db 246 TCGCTGAGTCCCTCGGCTTTACGGGCGCGCGTTCATCATCACGGAGATCGACCCCATC 305
 QY 901 tgcgtctccaggtccatggaaggtaccaggtccgcgcgcacatcgaggaagtgtcaag 960
 Db 306 AAGCCCTCCAGCTCGGATGGCGGCTACGAGTACGACCACTGAGGAGGCTCGGCC 365
 QY 961 gatgtcatatctctgtatcacaggaactgcgcatatcatctctgtgtgacatgatg 1020
 Db 366 CGCGGTAAAGTCTCTGTTACCAACCGCGGTAAACCGGACATCATCTACCTGGCAAGCACTTC 425

QY 1021 gccacatgaagataaggctattgtcggttaacatcgccacttcgataacgaaattgat 1080
 Db 426 GAGTCTATCGCGGAGGATGATGCTGTCCAACTCGGCCACTT-----CGNNNNNNNN 479
 QY 1081 acagatggcctcatgaaatataccagggcatcaagcacatcccccaatcaagccagaatacagc 1140
 Db 480 NNN 539
 QY 1141 atgtgggaattccacagatggccagcgtatctctcttctgtgagggcgccgtctttaa 1200
 Db 540 CGCTTCACCATGAAGAACCGCCGCACATCATCTCTCGCTGAGGGCGCTCTCGTCAAC 599
 QY 1201 ctgggtcgctacaggtcaccatcttctgtatgtaacatgctcattcacaaacagaca 1260
 Db 600 CTCGGCTGTGCCACCGGCCACCGCTGCTGTATGCTCTCTTCTTCAACCCAGAGC 659
 QY 1261 ctgcgtcagctcagcctctctacgaaagagaggaatctcgagaagaagtcttaccact 1320
 Db 660 CTCGGCAGATCGCTCTGGACCAACCGCGAGAGATTCCCTCTGGGTGCACATGCTC 719
 QY 1321 ccgaagcatctcgatgaagaagtcgctcgcctccacctcgatctctctgtatgtccactt 1380
 Db 720 CGAAGGAGCTCGACGAGGAGTCCGCGCCGACCTTGGCAGCTCAACGTCAGGCTC 779
 QY 1381 acaagcttacacagagcaggtgctgactacatcaacgttcaggtgagggctcttacaag 1440
 Db 780 ACCACGCTCTCCGACGAGCAGTCCAAAGTACTTCGACATCCGGTCAACGGCCCGTACAAG 839
 QY 1441 tctgattcttaccgttatta 1460
 Db 840 CCGTCGCACACTACCGCTACTA 859

RESULT 9
 LOCUS BF263419 908 bp mRNA linear EST 23-OCT-2001
 DEFINITION HV_CEA0006L04f Hordeum vulgare seedling green leaf EST library
 HV_CEA0006L04f (Blumeria challenged) Hordeum vulgare cDNA clone
 BF263419
 ACCESSION BF263419.2 GI:13260807
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare
 Barley.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 908)
 REFERENCE Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D.,
 Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi
 D.W., Fenton, R.D., Oates, R. and Main, D.
 TITLE Development of a genetically and physically anchored EST resource
 for barley genomics: Blumeria infected and physically anchored EST resource
 seedling leaf cDNA library
 JOURNAL Unpublished (2001)
 COMMENT On Nov 17, 2000 this sequence version replaced gi:11194413.

Contact: Wing RA
 Clemson University
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 548
 Seq primer: AATTAACCTCTCACTAAAGG
 High quality sequence stop: 759.

FEATURES
 Source Location/Qualifiers
 1..908
 /organism="Hordeum vulgare"
 /cultivar="C116155 (M1a13)"
 /db_xref="taxon:4513"
 /clone="HV_CEA0006L04f"
 /clone_lib="Hordeum vulgare seedling green leaf EST

library HVCDNA0004 (Blumeria challenged)
/tissue_type="seedling green leaf"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
C.I. 16155 (Mla13) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AvrMla13
) of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the TJ Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
cDNA library was made, and 1 million pfu were in vivo
excised to give Bluescript SK(-) cDNA phagemids (Choi,
Close). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT 205 a 276 c 234 g 193 t

Query Match 21.6%; Score 316.2; DB 10; Length 908;
Best Local Similarity 65.9%; Pred. No. 1e-72;
Matches 475; Conservative 0; Mismatches 243; Indels 3; Gaps 1;

QY 540 cgtcttctacacacagcaggtcttccacgaagacacacactggcacacagttgc 599
DB 9 CGTCTCACCATCATCCGGCAGCGCTCAAGACCGCGCCAGCAAGTACCGCAAGATGAA 68
QY 600 tgcgggcatgaacggtgtttccgaagagacacacacaggtgtccacgcctctaccagct 659
DB 69 GGAGAGGCTCGTCGGTGTCTCCGAGGAGACACACGCGGCTCAAGAGGCTCTACCCAGAT 128
QY 660 cgagaaggaggcaaacctctctccagccatcaacgtcaacgacgctgttcaaaagtc 719
DB 129 GCAGAGGTCGGCAGCCCTCTCTTCCCGCCATCAACGCTCAACGACCTCCGTCACCAAGAG 188
QY 720 caagtcgataacatctacggtgcgcgcacactcccttatcgatggtatcaacccgtgttc 779
DB 189 CAAGTTTGACACACTTACGGTTCGGTTCCTCTCTCTCCCTGACGGTCTCATGAGGGCCAC 248
QY 780 cgtatgtatgatcgccggaacagacagctctcgtcatgggttaagcgatgtcgcaagg 839
DB 249 TGATGTTATGATTGCGGGTAAGGTCCCGTGGTCTCGGTTATGTTGGTGTGGCAAGGG 308
QY 840 ctgcgctcaatccctccgtggccaaagcgtcgtgttatcatcacagaactccgacccaat 899
DB 309 CTGTGCGCGCGGACATCAACGAGGCTGGTGGCCGTGTGATCGTGCACAGATCGACCCCAT 368
QY 900 ctgcgctctccagggtgccatgaaggtaccagggtccgcgcgcatcaggaagtcgtaaa 959
DB 369 CTGTGCTCTTACGGCCCTGATGAGGGGCAATCCAGATCTCTCACCCTTGGAGGATGTGTCTC 428
QY 960 ggaatgcatactctgttacatgcacagaaactcgatctatctctgttgatgatg 1019
DB 429 TGAGGCTGATATTTTGTGACCAACACCGGAACAAAGACATCATCATGTTGGACCAAT 488
QY 1020 gggccagatgaaggatgaaggtattgtcgttaacatcgccacttcgataacgaattga 1079

Db 489 GAGGAAGATGAAGAACAAACGCGCATTTGTCGCAACATTTGGTCTACTTTGACAAACGAGATCGA 548
QY 1080 tacagatggcctcatgaataaccaggcatcaagcacatcccaatcaagccagaatacga 1139
Db 549 CATGAACGGCTTGGAGACCTACCCGGTGTCAAGCGCATCAACCATCAAGCCCAACTGA 608
QY 1140 catgtgggaattccacagatggcca---cgatatccctccctcttctgctgagggcgcccttct 1196
Db 609 CCGTGGGTCTTCCCTGAAACCAAGACCGGCATCATTTGTTCTTCTGCTGAGGTGCTCTGAT 668
QY 1197 taaccttgctgcgtacaggtcacccatcttctcgttatgtcgaatgcatcacaaacca 1256
Db 669 GAACCTGGATGTGCACATGGCCACCCCACTTTGTCATGCTCTGCTCTTCTCACTAACCA 728
QY 1257 g 1257
Db 729 G 729

RESULT 10
BF253807
LOCUS
DEFINITION
BF253807 867 bp mRNA linear EST 22-OCT-2001
HVSMEF0002C21f Hordeum vulgare seedling root EST library HVCDNA0007
(Etiolated and unstressed) Hordeum vulgare cDNA clone
HVSMEF0002C21f, mRNA sequence.
ACCESSION
VERSION BF253807.2 GI:13116738
KEYWORDS
SOURCE EST.
ORGANISM
barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE
AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
R.D., Oates, R. and Main, D.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics: Morex unstressed seedling root cDNA library
Unpublished (2001)
JOURNAL On Nov 16, 2000 this sequence version replaced g1:11182912.
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAAACCTCACTAAAGGG
High quality sequence stop: 730.
FEATURES
source
1..867
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEF0002C21f"
/clone_lib="Hordeum vulgare seedling root EST library
HVCDNA0007 (Etiolated and unstressed)"
/tissue_type="Seedling root"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedling roots were
then harvested, total RNA was prepared, poly(A) RNA was
purified, one primary unamplified cDNA library was made,
and 1 million pfu were in vivo excised to give Bluescript
SK(-) cDNA phagemids. These steps were performed in the TJ
Close laboratory at the University of California,
Riverside (Choi, Close, Fenton). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)

(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see close TJ, Wing R, Kleinohs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)

BASE COUNT	192 a	257 c	226 g	190 t	2 others
ORIGIN					
Query Match	21.5%; Score 313.6; DB 10; Length 867;				
Best Local Similarity	65.3%; Pred No. 4.9e-72;				
Matches	509; Conservative	0; Mismatches	261; Indels	10; Gaps	3;
QY	586 cagccatcaacgctcagcagcgtgttacaaagtcgaagttcgataacatctacgctgccc 745				
Db	1 CCGCATAAACCCCAACGACTCCGTCACCAAGAGCAAGTTTGACACCTTTACGGTTGCC 60				
QY	746 gccactcccttatcgatgtatcaaccggtgtccgatgtcatgctgacgagcagag 805				
Db	61 GTCACTCTCTCCCTGACCGGTCTCATGAGGGCCACTGATGTATGATCGCCGGTAAGTCG 120				
QY	806 ctctcgtcatgttgatggcgttcgcaagggctgcgctcaatccctccgctggccaa 865				
Db	121 CCGTGGTCTCGGTTATGCTGATGTGGCAAGGGCTGTGCGCCGCACTCAACGAGCTG 180				
QY	866 qcgtctcggttatcatcaacagactcgaccacaaatctgcgtctccaggtgcgcatgga 925				
Db	181 GTGCCGGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 240				
QY	926 gctaccaggtccgcgcgtcaggaagtcgcaaggtgcgctcaatccctccgctggcc 1045				
Db	241 GCATCAGATCTCTACCTTGGAGGATGTCTCTGAGGCTGATATCTTTGTGACCACCA 300				
QY	986 caggaactcgatcatctctctgttgatgatgtgcccagatgaaggatgaagctatt 1045				
Db	301 CCGGAACAAGGACATCATCATGTGTGACCATGAGGAAGATGAAGAACAACGCCATTG 360				
QY	1046 tcggttaacatcgccacttcgatacaaaatgatcacagatggcctcatgaataccca 1105				
Db	361 TCTGCAACATGGTCACTTTGACACGAGATCGACATGAACGGCTTTGAGACTACCCCG 420				
QY	1106 gcatcaagcactcccaatcaagcagatacgcacatgtgggaattcccagatggcca 1163				
Db	421 GTGTCATGCGCATCACCATCAAGCCCAAGACTGACCGTGGCTTCCCTTGACCAAGA 480				
QY	1164 -cgctatccctctctgtgagggcgccctctttaaacttggtgcgtacaggtcacc 1222				
Db	481 CCGGCAATCATGTCTTGTGAGGGTGTCTGTGATGAACCTTGGATGTGCACTGGCCACC 540				
QY	1223 catcttcgttatgtcaatgctcattcaaaacagacactcgctcagctcgacctta 1282				
Db	541 CCAGCTTGTGTCATGCTCTCTTCTACTAACCAAGTTATTGCCAGCTCGAGTTGTGG 600				
QY	1283 aaaa-----agaggaatctcagagaagaggtttac-acacttcccgaagatctc 1335				
Db	601 ACAGAAAGGCCAGCGGNCAGTACGAGAAGAGGTGTAGTTTCTCCCAAGCACTCGAC 660				
QY	1336 gaagaagtcgctcgctccacactcgatctcgtatgtccacttacaagaattcacag 1395				
Db	661 GAGAAAGTTGCGGCCCTTCCACTTGGGCAAGCTCGCGCGAGGCTGACCAAGTAACCA 720				
QY	1396 aagcaggtgactacatcaacgttccagttgagggctctctacagtgctgactacagt 1455				
Db	721 TCCCACTCTGACTACATTACCATCCGACCGAGGGTCCCTAAAGCCGCGCCCTTACCGT 780				

RESULT 11
BE559309
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BE559309 809 bp mRNA linear EST 23-OCT-2001
HV_CEB0023J02f Hordeum vulgare seedling green leaf EST library
HVCNDA0005 (Blumeria challenged) Hordeum vulgare cDNA clone
HV_CEB0023J02f, mRNA sequence.
BE559309 2 GI:13266539
EST.
barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 809)
Wing, R., Close, T.J., Kleinohs, A., Wise, R., Wei, F., Begum, D.,
Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi,
D.W., Fenton, R.D., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected incompatible (Mla6) seedling
leaf cDNA library
Unpublished (2001)
On Aug 14, 2000 this sequence version replaced gi:9823715.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 459
Seq primer: AATTACCTCTCACTAAAGGG
High quality sequence stop: 729.
Location/Qualifiers
1..809
/organism="Hordeum vulgare"
/cultivar="C116151 (Mla6)"
/db_xref="taxon:4513"
/clone="HV_CEB0023J02f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCNDA0005 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="SOLR"
/note="Vector: lambdaZAP; Site 1: EOR1; Site 2: Xho1;
C.I. 16151 (Mla6) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate 5874 (AvrMla6
) of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the TJ Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagemids (Choi, Close). Phagemids were plated and picked
at the Clemson University Genomics Institute (CUGI) (Begum
, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
this clone see <http://www.genome.clemson.edu/orders> Also
see close TJ, Wing R, Kleinohs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT
189 a 234 c 204 g 181 t 1 others

ORIGIN

Query Match 21.3%; Score 311; DB 10; Length 809;

Best Local Similarity 64.8%; Pred. No. 2.3e-71;

Matches 496; Conservative 0; Mismatches 261; Indels 9; Gaps 2;

QY 698 tcaacgacgtgtttacaaagtcacaggttcgataaatactatcaggtgcgcacccctccta 757
 DB 1 TCAACGACTCGTCCACCAAGACGAAGTTTGACAACTTACGGTTGCGCTCCTCTCC 60
 QY 758 tcatgtatcaacgctgtcccgatgtcatgtatgcggcggaagacagctctcgtcatgg 817
 DB 61 CTGACGCTCATGAGGCCACTGATGTATGATTTGCCGTAAGTTCGGCTGTGCG 120
 QY 818 gttacggcgatgtcggcaagggctgcgtcaatccctcgtgcccgaagcgctcgcgtta 877
 DB 121 GTTATGCTGATGTTGGCAAGGGCTTGCGCGCGCACTCAAGCAGGCTGTGCGCGTGA 180
 QY 878 tcatcagaactgcagccaatctgcctctccaggtcccatggaaggtaccaggtcc 937
 DB 181 TCGTGACAGAGATGACCCCACTCTGTCTTCAGGCCCTGTATGGAGGCACTCAGATCC 240
 QY 938 gcccatcgaggaagtcgaaggtgcgtatctcttcgttaccatgcacaggaactgcg 997
 DB 241 TCACCTTGGAGGATGTTCTCTGAGGCTGATATCTTGTGACCAACCCGCGAAACAAG 300
 QY 998 atatcatctgttgacatgatgcccagatgaaggaataagctattgtcggtaacatcg 1057
 DB 301 ACATCATCATGTGGTGCACCATGAGGAAGATGAAGAACAACGCCATGTGTGCAACATG 360
 QY 1058 gccacttcgaacgaataatgatacagatggcctctgaaataccacaggtcaacagaca 1117
 DB 361 CTCACCTTGACACAGATCGACATGAACGGCTTGAGACCTTACCCGGGTGCAAGCGCA 420
 QY 1118 tcccaatcaagccagatcatcatgttgggaattcccgatggcca---cgctatctcc 1174
 DB 421 TCACCATCAAGCCCGACAGTGAACGGCTGGGTCTTCCTGAGACCAAGACCGGCATCATG 480
 QY 1175 ttctgtcagggcgccctcttaaccttggctgcgtacaggtcacccatcttctgta 1234
 DB 481 TTCCTGCTCTTCTACCTAACCAAGTATGATGATGATGATGATGATGATGATGATGATG 540
 QY 1235 tgcattgtcattcaacacacagacactgcgtcagctcagctcagctcagctcagctcagct 1290
 DB 541 TGCTCTGCTCTTCTACCTAACCAAGTATGATGATGATGATGATGATGATGATGATGATG 600
 QY 1291 --ggaatctcaggaaggtttacacacttcagacatcgcgtatcaggaagtcgctc 1348
 DB 601 CCGGCCAGTACCAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 1349 gctccacctcggatctcgtatgtccacttacaagcttacaagcttacaagcaggtcact 1408
 DB 661 GCCTTCACCTTGGCAAACTCGGGCGCCAGCTGACCAAACTTAACAAATCCCAATCTGACT 720
 QY 1409 acatacaactccaggtgaggttccttacaagctcagctcagctcagctcagctcagct 1454
 DB 721 ACCTTAACATCCGATCGAGGGGCGCTTACAAGCGCGCGGCTACCGG 766

RESULT 12
 LOCUS BM411809 804 bp mRNA linear EST 22-JAN-2002
 DEFINITION EST586125 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLE57P14 5' end, mRNA sequence.
 ACCESSION BM411809
 VERSION BM411809.1 GI:18263428
 KEYWORDS EST
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Lycopersicon.
 1 (bases 1 to 804)
 Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai
 , J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning
 , C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 Unpublished (2002)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3.

FEATURES
 source

Location/Qualifiers
 1. 804
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone_lib="tomato breaker fruit"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLs"
 /note="Vector: pBluescriptSKmCuadapt; site:1: EcoRI;
 Site:2: XhoI; supplier: Boyce Thompson Institute; Fruit
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 lycopene accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."

BASE COUNT
 ORIGIN

207 a 174 c 196 g 227 t

Query Match 21.0%; Score 307.4; DB 10; Length 804;
 Best Local Similarity 63.0%; Pred. No. 2e-70;
 Matches 492; Conservative 0; Mismatches 286; Indels 3; Gaps 1;
 QY 502 ccagagccacagaagctgacacacctcgaataccgctgcgtctctgtctacactcaagag 561
 DB 1 CCAGATCCCACTTCTACTGACAAATGTTGAGTTTCAACTTGTGCTTACTATTATTAAAGAG 60
 QY 562 gtcttcaaccaagagaagaccactggcacacagttgctgcggcatgaacggtgttcc 621
 DB 61 AGCTTAAAGACTGATCCATTAAAGTACATTAAGATGAAGGAGAGACTTGTGTGTCT 120
 QY 622 gaagagacaacaacaggtgtccaccgctctaccagctcgagaagaggaggaacacccctc 681
 DB 121 GAGGAAACTACCACCTGCTGTGTAAGAGGCTTTACCAATGCAGGCTAATGGATCTTTGCTT 180
 QY 682 ttcccgccatacaagctcaacgacgctgttacaaggtccaagttcgataacacatcacgc 741
 DB 181 TTTCTCTGCTATCAATGTTAATGACTCTGTACCAAGACGAAGTTTGACAACTTGTATGA 240
 QY 742 tgcgcacactcccttctcgtatgtatcaacccgtgcttccgctgctcgtatgcagtcgagcaag 801
 DB 241 TGCAGCCACTCACTTCCCGATGGTCTCATGAGGGCTACTGATGTTATGATGTTGTTGAAG 300
 QY 802 acagctctcgtatcgttgcaggtgcgtcgcgaagggctgcgtcctcctccctccgtg 861
 DB 301 GTTGCTCTTGTGCTGTTGAGATGTCGCAAGGATGTGCTGCTGCCATGAACAA 360
 QY 862 caagcgctcgcgttaccatcacagaactcgacccaacttcgctctccaggtcgcagt 921
 DB 361 GCTGGTCCCGCTGTGATGTGCTACTGAGATTGACCAACTGTGCTCTCCAGGCTACCATG 420
 QY 922 gaaggtaccaggtcccgccgc 981
 DB 421 GAAGGCTCCAGGTTCTTCCCTCTTGGAGATGTTGTTCTGTAGGTTGATATCTTTGTGACC 480
 QY 982 tgcacaggaactcgcgtatcatctctgttgacatgatggtcccgatgagggatgaaggct 1041

```

Db 481 ACACCGGTAAACAGGACATCATCATGTTGACCAATGAGGAAGATGAAGAACAATGCC 540
QY 1042 attgctggttaacatcgccacttcgataacacgaattgatacagatggccctcagaataac 1101
Db 541 ATTGCTGCAACATTTGGTCACTTTGACAAAGAAATCGACATGCATGCTTGTGAACCTTC 600
QY 1102 ccaggcaataagacacatcccaataaagccagaataacgacatgtgggaattccagatggc 1161
Db 601 CCRGGTGTGAAGAGGATACAAATCAAGCCACAAACCGACAGATGGTCTTCCAGACACC 660
QY 1162 cac---gctatccctcttctgctgagggcgccctcttaacctgttggaattccagatggc 1218
Db 661 AACAGTGGCATCATGCTTGGCCGAGGGTGTCTCATGAACCTTGGGATGTGCCACTGGA 720
QY 1219 caccatcttctgattgataatgatacacaacagacactgcctcagctcgacctc 1278
Db 721 CACCCCAAGTTTGTGTGATGCTTGTCTTCTTCTACCAAGTCATTTGCCCACTCAAGTTG 780
QY 1279 t 1279
Db 781 T 781

RESULT 13
BG444868
LOCUS
DEFINITION
GA_Ea0025010f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0025010f, mRNA sequence.
ACCESSION
BG444868
VERSION
BG444868.1 GI:13354520
KEYWORDS
EST.
SOURCE
Gossypium arboreum.
ORGANISM
Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE
1 (bases 1 to 916)
AUTHORS
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL
Unpublished (2000)
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Email: rwing@clemson.edu
Fax: 864 656 4293
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
source
1..916
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0025010f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT
225 a 209 c 264 g 218 t
ORIGIN

Query Match 20.8%; Score 303.4; DB 10; Length 916;
Best Local Similarity 63.0%; Pred. No. 2.5e-69;
Matches 521; Conservative 0; Mismatches 296; Indels 10; Gaps 3;
QY 625 gagacaacaaggtgtccaccgctctaccagctcagagaggggcaactctcttc 684

```

```

Db 12 GAAATACAAACGGTGTAAAGAGCTTTATCAGATGCAGGCCAATGGAACCTTGTGTTC 71
QY 685 ccagccatcaacgctcaacgacgctgttcaaaagtcacagttccgataacatctacggctgc 744
Db 72 CCTGCTATTAATGTCAATGACTCTGTCAACCAAGAGCAAGTTCGATAACTTGTATGCGATCC 131
QY 745 cgcacatccctctatcgatggatcaaccgctgtcccgatgtcatgtatcggtgggcaagaca 804
Db 132 CBTCACTCTCTTCCCGATGGTGTGATGAGAGCTACCCATGTGATGCTGCTGCTTGAAGGTC 191
QY 805 gctctcgctcatgggttacggcgatgtcgccaagggtcgctcaatccctccgtggccaa 864
Db 192 GCTGTGTCTGTGGTTATGTTGATGTCGCAAGGGTGTGCTGCTGCTTGAAGCAAGCC 251
QY 865 ggcgtcgcgttatcatcatcacagaactcgaccgaatctgcctccaggtgcatatgaa 924
Db 252 GGTGCTCCGCTCATTTGTCACCGAGATTGATCCCATGTCGCCCTTCAGGCTCTCATGAA 311
QY 925 ggcataccaggtccgcgcatcgaggaagtcgcaagatgtcgatctcgatatcttcgttacatgc 984
Db 312 GGACTTCAAGTTTTCACCTTGAGGATGTTGTCGAGGCTGATATCTTTGTCCACCAGC 371
QY 985 acaggaactgcgatcatctctgttgacatgatggccagatgaaggataaaggctatt 1044
Db 372 ACTGTTAAACAAGGACATCATCATGTTTAAACCATGAGGAAGATGAAGAACAATGCCATTT 431
QY 1045 gtcgtaacatcgccacttcgataacgaattgatcacagatggcctcatgaataacca 1104
Db 432 GTTTGCAACATCGGTCACTTCGACAAATGAATTGACATGCTCGCTCTCGAGAAATATCTCT 491
QY 1105 ggcatacagcacatcccaatcaagccagatacgacatgtgggaattccagatggcca - 1163
Db 492 GGTGCTCAAGCGCATTACCATCAAGCTCAACCGGATAGGTGGTCTTCCCTGAACCAAC 551
QY 1164 --cgctatctctctctgtgagggcgccctctttaaacttggtgcgctacaggtcac 1221
Db 552 ACCGGCATCGTGTGTGGCTGAAGAGCTCTCATGAACCTTGGTGTGCTACTGGCCAC 611
QY 1222 ccattcttcgttatgtaaatgtcattcacaacagacactcgctcagctcgacctetac 1281
Db 612 CTTAGCTTTGGCATGTCTATGCTGCTTCACTAACCCAGGTGATCGCCAGCTGGAGCTGTGG 671
QY 1282 gaaag-----agagaaatctcgagaagaaggtttacacacttcggaagcatctcgat 1335
Db 672 TAAGAGAATGCAACAGCAAGTGGGAGAGAAAGAGGGGGGTTTGCTTAACACCTTGAT 731
QY 1336 gaagaagtcgctgcctccactcggatctctcgatgtccacttacaagcttacacag 1395
Db 732 GAGAAAGTGGTGTCTTGTAGCTTTGAAAGCT-GAGGGGTACCTACCAAGCTCACACAGA 790
QY 1396 aagcaggctgaactacatcaacgcttccagttgaggtccttacaagtc 1442
Db 791 GATCAAGACAGCTAAATTAAGCGGGCCCTAGGGGAGCTTAAAGGCC 837

RESULT 14
BF863943
LOCUS
DEFINITION
769 bp mRNA linear EST 19-JAN-2001
963047G08.y1 C. reinhardtii CC-1690, Stress condition I, normalized
, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION
BF863943
VERSION
BF863943.1 GI:12254087
KEYWORDS
EST.
SOURCE
Chlamydomonas reinhardtii.
ORGANISM
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE
1 (bases 1 to 769)
AUTHORS
Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,
Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in

```


JOURNAL COMMENT

Vascular plants; project phase 3
Unpublished (2000)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177

FEATURES

1. 769
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21g"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
/note="Vector: pBluescript
XhoI: This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr),
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr,
4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
Zap clones by superinfection with ExAssist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

BASE COUNT	170 a	256 c	221 g	122 t
ORIGIN				

Query Match	20.7%	Score 302.8;	DB 10;	Length 769;
Best Local Similarity	67.4%	Pred. No. 3.3e-69;		
Matches 427;	Conservative	0;	Mismatches 207;	Indels 0;
Gaps 0;				

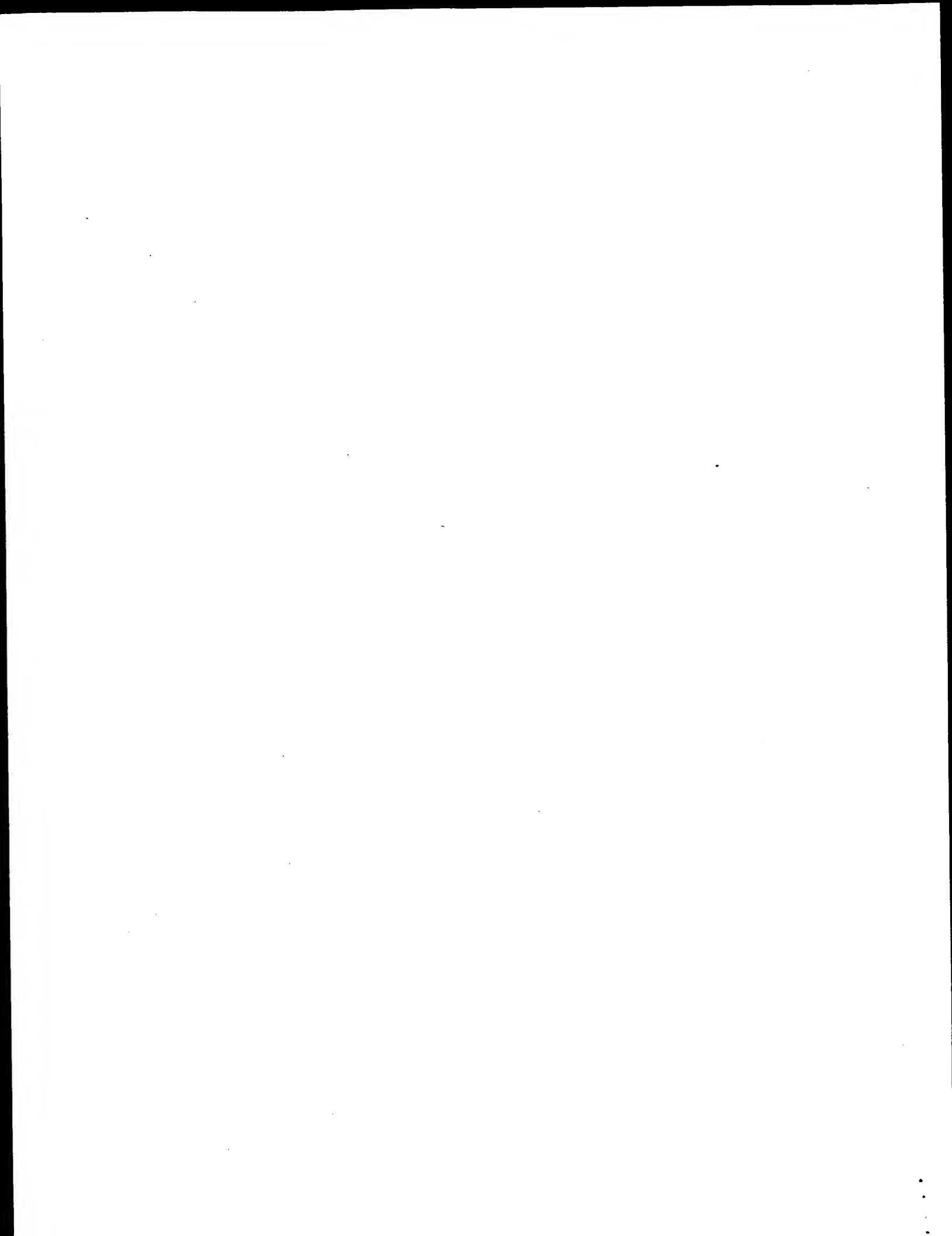
556 agctcgagaaggagggaactctctccagccatacaactcaacgacgctgttaca 715
Db AGATGACAGGTAACGGCAGCGTGTGTTCCCGCCATCAACGTGACGACTCGGTACCA 186
716 agtccaaagtccgaataacatcacggtcgcgcgcacatcccttatcgattatcaacacgtg 775
Db AGTCCAAGTTCGACAACGTTACGGCTGCGCGCACTTCGCTGCCGACGGCATCATGGCG 246
776 ctccgatgtcatgatcggcgcgaagacagactctcgtcatggtttacggcgatctcgca 835
Db CCAGTACGCTGATGATTCCCGGCAAGACCGCTTATCGCCGCTACCGTGTGATGTGGCA 306
836 agggctcgctcaatccctccctggtgccaaaggcgtcgcttatcatcacagaactgacc 895
Db AGGGCTCGCCAGCGCCATGAAGCGCGCGGTGCCGCAACCATTTGTCGGAGATCGACC 366
896 caatctcgctctccaggtcgccatggaaagctaccaggtccgcgcgcgaatcgagaagtgc 955
Db CCATCTCGCGCTCGAGCGGCCATTTGAGGGGTACCAAGGTCCAGCCCCCTGGAGACTGCC 426
956 tcaaggatgtcgatatctcgtttacatgcacaggaactcgcgatatactctctgtgaca 1015
Db TGGACACCGCGACATCTTCATCACCAACCCGGCAACAGGACATCATGTGGTGTGAGC 486
1016 tgatggcccagataaaggataaggctattgttcggttaacatcgcccaacttcgataacgaaa 1075
Db ACATGGCCAAAGATGAAGAACAAACGCCATTGTGGGCAACATTTGGCCACTTCGACAAACGAGG 546
1076 ttgatacagatggcctcatgaaatacccaggatcatgaagcacatcccatacagccagaat 1135
Db TGGACATGGCGCGCTGTACGCTTGCCCGGCATCAAGCGCCAGAACATCAAGCCCCCAGG 606
1136 acgacatgtgggaattcccagatggccaagctatcctcctcttcttgctgagggccgacttc 1195
Db TGGACCGCTTCATCTTCCCGACGACACGGCGGTGATTGATCTGATGATGGGCCGACATGTC 666

see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>.)

	BASE COUNT	ORIGIN
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	0	0
39	0	0
40	0	0
41	0	0
42	0	0
43	0	0
44	0	0
45	0	0
46	0	0
47	0	0
48	0	0
49	0	0
50	0	0
51	0	0
52	0	0
53	0	0
54	0	0
55	0	0
56	0	0
57	0	0
58	0	0
59	0	0
60	0	0
61	0	0
62	0	0
63	0	0
64	0	0
65	0	0
66	0	0
67	0	0
68	0	0
69	0	0
70	0	0
71	0	0
72	0	0
73	0	0
74	0	0
75	0	0
76	0	0
77	0	0
78	0	0
79	0	0
80	0	0
81	0	0
82	0	0
83	0	0
84	0	0
85	0	0
86	0	0
87	0	0
88	0	0
89	0	0
90	0	0
91	0	0
92	0	0
93	0	0
94	0	0
95	0	0
96	0	0
97	0	0
98	0	0
99	0	0
100	0	0

Query Match	20.6%	Score 301.2	DB 10	Length 844	
Best Local Similarity	64.5%	Pred. No. 9.2e-69			
Matches 485	Conservative	0	Mismatches 258	Indels 9	Gaps 2
QY 702	cgacgctgtttacaaagtc	ccaagtttcgaataacatctacagctcgccgacatcccttatcg	761		
DB					
DB					
DB					
QY 762	tggtatcaaccgtgcttccgat	ctcatgatcgcgcaagacacgctctctgctcatggtta	821		
DB					
DB					
DB					
QY 822	cgccgatgtcgccaaaggtcgctc	aatccctcogtggccaaagcgctcggttatcat	881		
DB					
DB					
DB					
QY 882	cacgaactcgaccctaactcgctct	ccaggctgccatgaagctaccaggctccgcgcg	941		
DB					
DB					
DB					
QY 942	catcgaggaagtcgtccaaaggtctgat	tcttctgtacatgcacaggaaactcgcatat	1001		
DB					
DB					
DB					
QY 1002	catctctgttcacatgatggccagatgaaggtataaggtatt	gtcgttaacatcggcca	1061		
DB					
DB					
DB					
QY 1062	cttcgatcaacgaattgatcacagatggcctcatgaataccaggcatcaacgacatccc	1121			
DB					
DB					
DB					
QY 1122	aatacagccgaatacagatgtgggaattcccatgtggcca	---cgctatctctctctct	1178		
DB					
DB					
DB					
QY 1179	tgctgagggcgccctcttaacttgctgcgtacaggtcacccatcttctggttatgtc	1238			
DB					
DB					
DB					
QY 1239	aatgtcattcacaacccagacactcgctcagctcgacctctacgaaag	-----agagg	1292		
DB					
DB					
DB					
QY 1293	aaatctcgagaagaagtttacacacttccgaagcatctcgatgaagaagtcgctcgct	1352			
DB					
DB					
DB					
QY 1353	ccacctggatctctcgatgtccaccttacaagcttacaagaagcaggtcgactacat	1412			
DB					
DB					
DB					
QY 1413	caacgttccagttgagggctcattacaagtctg	1444			
DB					
DB					
DB					
DB					

Search completed: August 17, 2002, 05:50:33
Job time: 14076 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:44:41 ; Search time 268.29 Seconds
(without alignments)
9349.629 Million cell updates/sec

Title: US-09-759-990-1

Perfect score: 1461

Sequence: 1 atggttgcaaatcaccctac.....ctgatgcttaccgtatttaa 1461

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1459.4	99.9	1461	22	Trichomonas vagina
2	1440.2	98.6	1461	22	Trichomonas vagina
3	596.4	40.8	1557	22	Corynebacterium gl
4	596.4	40.8	1557	23	C. glutamicum gene
5	596.4	40.8	349980	22	C glutamicum codin
6	594	40.7	1422	22	C glutamicum codin
7	535.4	36.6	1488	21	M. tuberculosis an
8	529.6	36.2	1396	22	Corynebacterium gl
9	529.6	36.2	1396	23	C. glutamicum gene

10	508.8	34.8	1865	21	AAC44037	Zea mays DNA fragm
11	506.6	34.7	1727	21	AAC46315	Arabidopsis thalia
12	497.8	34.1	1728	21	AAC39836	Arabidopsis thalia
13	489	33.5	1767	17	AAQ44513	Asparagus s-adenos
14	488.4	33.4	1812	13	AAQ29419	Nicotiana tabacum
15	488.4	33.4	1812	17	AAT23194	S-adenosylhomocyst
16	471.4	32.3	1494	21	AAC46817	Arabidopsis thalia
17	375.4	25.7	4411529	22	AAI99682	Mycobacterium tube
C	331	22.7	2200	21	AAC77823	Human cancer assoc
C	302	20.7	513445	22	AAI61373	Soybean 318013 reg
20	297.8	20.4	1941	17	AAT44515	Arabidopsis S-aden
21	293.4	20.1	1398	21	AAT98330	A. thaliana gene i
C	288.4	19.7	720	22	AAH63803	C glutamicum codin
C	278.6	19.1	708	22	AAH65802	C glutamicum codin
24	244	16.7	2563	19	AAV28617	Nucleotide sequenc
25	244	16.7	3616	20	AAV73924	Human SAHH DNA #1.
26	233.6	16.0	1566	23	ABL03693	Drosophila melanog
27	223.2	15.3	2211	22	AAC91228	Human S-adenosylho
C	219.2	12.3	5030	23	AA576216	DNA encoding novel
C	209	16.1	4406	23	ABL03692	Drosophila melanog
C	161.2	11.0	4406	23	ABL03692	Genomic fragment #
C	149.8	10.3	58909	22	AAF28543	Pyrococcus abyssi
C	146.8	10.0	349980	22	AAH41226	Ripening banana pu
32	137	9.4	793	19	AAV28651	Fusarium venenatum
C	125.2	8.6	1949	21	AAF07609	Human secreted pro
34	121	8.3	363	21	AAC01924	Human SAHH cDNA.
35	120.8	8.3	289	16	AAQ98682	Aspergillus oryzae
36	120.8	8.3	656	21	AAF14585	Plant microsateili
37	108.2	7.4	371	21	AAA31988	Ripening banana pu
38	108.2	7.4	722	19	AAV28650	Plant microsateili
39	106.8	7.3	374	21	AAA31802	Ripening banana pu
40	105	7.2	285	16	AAQ98681	Plant microsateili
41	99	6.8	3795	22	AAC87646	SAHH-GSE. Mus sp.
42	96	6.6	352	21	AAA31634	Brassica napus myo
C	95.4	6.5	573	22	AAI90254	Plant microsateili
C	91.4	6.3	1000	22	AAF91420	Human polynucleoti
45	87	6.0	4761	23	ABL14956	Moraxella catarrha
						Drosophila melanog

ALIGNMENTS

RESULT 1
ID AAS09278 standard; DNA; 1461 BP.

XX AAS09278;

XX 24-OCT-2001 (first entry)

DE Trichomonas vaginalis recombinant His-SAHH gene sequence.

XX SAHH; S-adenosyl homocysteinase; S-adenosyl homocysteine hydrolase;
KW S-adenosylmethionine; SAM; glycine N-methyltransferase; GMT; His-SAHH;
KW cancer; malaria; arthritis; SAH; mutant; ds.

XX Trichomonas vaginalis.

OS Synthetic.

XX	Key	Location/Qualifiers
FT	mutation	replace (19,G)
FT	mutation	/*tag= a
FT	mutation	replace (201,G)
FT	mutation	/*tag= b
FT	mutation	replace (207,T)
FT	mutation	/*tag= c
FT	mutation	replace (210,T)
FT	mutation	/*tag= d
FT	mutation	replace (501,C)
FT	mutation	/*tag= e
FT	mutation	replace (744,T)
FT	mutation	/*tag= f
FT	mutation	replace (834,G)
FT	mutation	/*tag= g

mutation replace (897,T)
 mutation /*tag= h
 mutation replace (917,T)
 mutation /*tag= i
 mutation replace (1206,T)
 mutation /*tag= j
 mutation replace (1304,T)
 mutation /*tag= k
 mutation replace (1346,T)
 mutation /*tag= l

XX WO200151651-A2.

PN XX

XX 19-JUL-2001.

XX 12-JAN-2001; 2001WO-US01114.

PF 14-JAN-2000; 2000US-0176444.

XX (ANTI-) ANTICANCER INC.

XX Hoffman RM, Xu M, Han Q;

XX WPI; 2001-451863/48.

XX Assessing therapeutic levels of S-adenosylmethionine comprises

PT measuring reaction products in sample containing glycine

PT N-methyltransferase, (His) S-adenosyl homocysteine hydrolase and

PT glycine

XX Claim 6; Fig 6A-6B; 28pp; English.

XX The present sequence for His-SAHH represents a novel recombinant gene

CC sequence for Trichomonas vaginalis S-adenosyl homocysteinase

CC (SAHH, S-adenosyl homocysteine hydrolase). The wild type SAHH gene

CC sequence (AA09277) is used to generate the His-SAHH gene sequence

CC which comprises various point mutations and an additional polynucleotide

CC sequence which encodes for an extra 6 N-terminal histidine residues.

CC The invention relates to a method for assessing therapeutic

CC levels of S-adenosylmethionine (SAM) in a biological fluid sample

CC comprising measuring one or more reaction products in a sample

CC containing glycine N-methyltransferase (GNMT), SAHH or His-SAHH, and

CC glycine, where the level of one or more products is directly

CC proportional to the level of SAM in the sample. The method is useful

CC for assaying therapeutic levels of SAM in a biological sample and may

CC be used as a part of a diagnostic protocol or as part of a therapeutic

CC protocol, where conditions or progress of the therapy may be monitored.

CC SAHH or His-SAHH may be used as a reagent, particularly in screening for

CC inhibitors and inactivators of the enzyme, for use as reagents themselves

CC as potential therapeutics, e.g. in cancer, malaria, arthritis and other

CC diseases. Recombinant SAHH may be used as a therapeutic cancer gene in

CC combination with S-adenosyl homocysteine (SAH) analogues.

XX Sequence 1461 BP; 372 A; 427 C; 335 G; 327 T; 0 other;

Query Match 99.9%; Score 1459.4; DB 22; Length 1461;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcttgcaaatcacctactagggtgctccattcgtagtacagaaatgccacatcaacctc 60

Db 1 atggcttgcaaatcacctactagggtgctccattcgtagtacagaaatgccacatcaacctc 60

QY 61 catgtctcgcgcgtaaggaaactacccttgcgtgagaagaaatgccaggtcttatggtt 120

Db 61 catgtctcgcgcgtaaggaaactacccttgcgtgagaagaaatgccaggtcttatggtt 120

QY 121 ctctgtagcggttatcttcgcttctaagcattgaagggtgtcagaatctctgttccctc 180

QY 121 ctctgtagcggttatcttcgcttctaagcattgaagggtgtcagaatctctgttccctc 180

Db 121 ctctgtagcggttatcttcgcttctaagcattgaagggtgtcagaatctctgttccctc 180

QY 181 cacatgacagtcacagacgcgtctctatcgagacactcacagctcttggtgatgctc 240

Db 181 cacatgacagtcacagacgcgtctctatcgagacactcacagctcttggtgatgctc 240

QY 241 agatgggtctctctcaacatctctctacacagatatacagccgctgctgtatctgtgctc 300

Db 241 agatgggtctctctcaacatctctctacacagatatacagccgctgctgtatctgtgctc 300

QY 301 ggcccaacagacacacacagagagccagccggtatccccagttcttcgctggaaggcgaa 360

Db 301 ggcccaacagacacacacagagagccagccggtatccccagttcttcgctggaaggcgaa 360

QY 361 acactccagaataactgggagaaacacataccgctctctcacatggccagatggtcaaggc 420

Db 361 acactccagaataactgggagaaacacataccgctctctcacatggccagatggtcaaggc 420

QY 421 ccacagcaggttgcgtgatggtggtgatgtctacactctctctccaaagggtctcgaa 480

Db 421 ccacagcaggttgcgtgatggtggtgatgtctacactctctctccaaagggtctcgaa 480

QY 481 ttccaaacagccggtgctgttccagagcccaacagagctgacaaactctgaataaccgtgc 540

Db 481 ttccaaacagccggtgctgttccagagcccaacagagctgacaaactctgaataaccgtgc 540

QY 541 gttcttgctacactcaagcaggtcttccaaacagagacagaaacacacacagttgct 600

Db 541 gttcttgctacactcaagcaggtcttccaaacagagacagaaacacacacagttgct 600

QY 601 gccggcatgaacggtgttctccgaagagacaaacacaggtgtccaccgctctaccagctc 660

Db 601 gccggcatgaacggtgttctccgaagagacaaacacaggtgtccaccgctctaccagctc 660

QY 661 gagaagggggcaaaactctctccagccatacaacgctcaaacgacgctgtataaagtcc 720

Db 661 gagaagggggcaaaactctctccagccatacaacgctcaaacgacgctgtataaagtcc 720

QY 721 aagttcgataacactagcgtgcgcgcactccctctatcgatggtatacccgctgctcc 780

Db 721 aagttcgataacactagcgtgcgcgcactccctctatcgatggtatacccgctgctcc 780

QY 781 gatgtcatgatcggcggcaagacagctctcgtatgggttacggcgagtgctggcaaggc 840

Db 781 gatgtcatgatcggcggcaagacagctctcgtatgggttacggcgagtgctggcaaggc 840

QY 841 tgcgtcaatccctcctcgtggcgaaggcgtcgcgttatctatcaacagaaactcgacccaatc 900

Db 841 tgcgtcaatccctcctcgtggcgaaggcgtcgcgttatctatcaacagaaactcgacccaatc 900

QY 901 tgcgtctccaggtgcctatgaaggctaccaggtccgcgcctcagaggaagtgcgtaag 960

Db 901 tgcgtctccaggtgcctatgaaggctaccaggtccgcgcctcagaggaagtgcgtaag 960

QY 961 gatgtcgatatcttcgtttacatgcacaggaactcgatatactctctgttgacatgatg 1020

Db 961 gatgtcgatatcttcgtttacatgcacaggaactcgatatactctctgttgacatgatg 1020

QY 1021 gccacagatgaagataaaggctattgtcggtataacatggccacttcgataacgaattgat 1080

Db 1021 gccacagatgaagataaaggctattgtcggtataacatggccacttcgataacgaattgat 1080

QY 1081 acagatggcctcatgaaatccccaggtcacaacacacacacacacacacacacacacacac 1140

Db 1081 acagatggcctcatgaaatccccaggtcacaacacacacacacacacacacacacacacac 1140

QY 1141 atgtgggaattccagatggccacgctatctctctctctctctctctctctctctctcttaac 1200

Db 1141 atgtgggaattccagatggccacgctatctctctctctctctctctctctctctcttaac 1200

QY 1201 ctgtgctgcgtacaggtcacccatctctctctctctctctctctctctctctcttaac 1260

Db 1201 ctgtgctgcgtacaggtcacccatctctctctctctctctctctctctctctcttaac 1260

QY 1261 ctgcgtcagctcagcctctacgaaagagaggaatctcgagaagaagggtttacacatt 1320

Db 1261 ctgcgtcagctcagcctctacgaaagagaggaatctcgagaagaagggtttacacatt 1320

[illegible]

Db	1016	gcaaaagaacatc	atcgtctcgcag	acattaaagcgat	gtagaagacacacgcg	actctcgg	1075	
Qy	1049	gtaacatcgccact	ctcgataa	cgaattgatacaga	tgtgcctcatgataa	taccagagca	1108	
Db	1076	gaatatcgccact	ctcgacagagat	cgcatactg	gcggtgtcgatccgcggcgga		1135	
Qy	1109	tcaagcacatccca	tccaatcaagc	gaataagaa	atgfyggaattccca	gagccacgcta	1168	
Db	1136	caagcgatcaaac	ctcaagcctcga	ggtctgaacctg	tgagaccttggc	gaaacagcgccgcgcga	1195	
Qy	1169	tccctctctctg	ctgagagcgccct	cttcaaccttg	ctgctgctatacag	gtccaccatctt	1228	
Db	1196	tcatcggtgctg	ccggagcggtcg	ctgctgagacctg	ggcaatgcca	ccggcgaacctctgt	1255	
Qy	1229	tctgtatgtcaat	gtatattcca	caaacacga	gacatctgtctg	gtcgtgactctctaa	gaaagaa	1288
Db	1256	tctgtatgtagca	aaacagctctgc	ctcaaacacga	gactgcgccga	atctgagactgtg	gacaaaga	1315
Qy	1269	gaggaatctcga	agaagaaggt	ttacaacattc	cgaaacatctcga	tgaatgaagaag	ctgc	1348
Db	1316	acgaagcagatca	gacaaacg	ggtgtaccgg	ctgtgcccaag	gacactcgacga	gagaaggtg	1375
Qy	1349	gctccacactcg	atctctcga	tgtcca	acttacaaggt	ttacacagaaga	gaggtgc	1400
Db	1376	gaatccatgtcg	agcccttgg	cggtgtacac	cttgacaagag	ctgtgacgaaga	gagagcggaat	1433
Qy	1409	acaatacgttcca	gtttgaggtcc	cttaacaagt	ctgtatgtctta	ccgtattatca	1461	
Db	1436	accttcggtcg	acgtctcga	ggtctcc	cttacaag	ccggagacatac	cgctatcga	1488

RESULT	8
AAAF71864	
ID	AAAF71864 standard; DNA; 1396 BP.
XX	
XX	AAAF71864;
AC	
XX	
DT	30-APR-2001 (first entry)
XX	
DE	Corynebacterium glutamicum MP protein nucleotide sequence SEQ ID NO:223
XX	
XX	Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW	fine chemical production; microorganism; organic acid; nucleoside;
KW	nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW	lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW	carbohydrate; aromatic compound; cofactor; polypeptide; enzyme; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	WO200100843-A2.
XX	
PD	04-JAN-2001.
XX	
PF	23-JUN-2000; 2000WO-IB00923.
XX	
PR	25-JUN-1999; 99US-0141031.
PR	01-JUL-1999; 99DE-1030476.
PR	02-JUL-1999; 99US-0142101.
PR	08-JUL-1999; 99DE-1031415.
PR	08-JUL-1999; 99DE-1031418.
PR	08-JUL-1999; 99DE-1031419.
PR	08-JUL-1999; 99DE-1031420.
PR	08-JUL-1999; 99DE-1031424.
PR	08-JUL-1999; 99DE-1031428.
PR	08-JUL-1999; 99DE-1031434.
PR	08-JUL-1999; 99DE-1031435.
PR	08-JUL-1999; 99DE-1031443.
PR	08-JUL-1999; 99DE-1031453.
PR	08-JUL-1999; 99DE-1031457.
PR	08-JUL-1999; 99DE-1031465.
PR	08-JUL-1999; 99DE-1031478.
PR	08-JUL-1999; 99DE-1031510.
PR	08-JUL-1999; 99DE-1031541.

PR 08-JUL-1999; 99DE-1031573.
PR 08-JUL-1999; 99DE-1031592.
PR 08-JUL-1999; 99DE-1031632.
PR 08-JUL-1999; 99DE-1031634.
PR 08-JUL-1999; 99DE-1031636.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032126.
PR 09-JUL-1999; 99DE-1032130.
PR 09-JUL-1999; 99DE-1032166.
PR 09-JUL-1999; 99DE-1032206.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 99DE-1032922.
PR 14-JUL-1999; 99DE-1032926.
PR 14-JUL-1999; 99DE-1032928.
PR 14-JUL-1999; 99DE-1033004.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 12-AUG-1999; 99US-0148613.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040832.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041380.
PR 31-AUG-1999; 99DE-1041394.
PR 31-AUG-1999; 99DE-1041396.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX
PA (BAD) BASF AG.
XX
PI Pompeius M, Kroegeer B, Schroeder H, Zelder O, Habernauer G;
XX
DR MPI; 2001-137957/14.
XX
DR P-PSDB; AAB79745.
XX
PT Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 3; Page 477-479; 1737pp; English.
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 1396 BP; 309 A; 396 C; 384 G; 307 T; 0 other;

Query Match	36.2%	Score	529.6;	DB	22;	Length	1396;
Best Local Similarity	64.3%	Pred. No.	4e-146;				
Matches	828;	Conservative	0;	Mismatches	454;	Indels	6;
				Gaps	2;		
QY	32	tcgagtaagaattccgcagatcaacccacccatggtcttcggcgcgaagaattacccttg	91				
DB	114	tggacttcaagaatgctgcagatcttctaactacagcagagcaggagctacccgattcgcttg	173				

Accession	Sequence	Position
QY 1172	tcctcttcgtgtagggccgcgcctcttaaccttgcgtcgtcgaagttccaccatcttcg	1231
Db 1248	tcgtccctcgtcgaagaagtcgcctggtttaaccttggcaacgcacccgcagaccocatcatctt	1307
QY 1232	ttaatgcacatgcatcttcacaaacagacactgcctaagctcgacactctaagaaagag	1291
Db 1308	tcattgtccaacctcttcgcgcgatacaagaaccttgcgcagatctgcgaactgttccaaaagag	1367
QY 1292	gaaalctcgagaagaaggtttaacact	1319
Db 1368	gacagtaacgagaacgagttctaccgct	1395

DJ	26-FEB-2002	(first entry)
XX		

c. glucanase gene #48 encoding metabolic pathway protein.

KW methionine biosynthesis pathway; lysine biosynthesis pathway;
 KW methionine biosynthesis pathway; large-scale production of fine chemical
 KW Corynebacterium diphtheriae; diphtheria; ds.
 KW

05 corynebacterium glutamicum.

PN WO200166573-A2
 YX

PD 13-SEP-2001.

22-DEC-2000; 2000WO-IB02035.

PR 09-MAR-2000; 2000US-187970P.
PB 23-JUN-2000 3000US-0605740.

PA (BADI) BASF AG.
 XY

Pi
Pompejus M., Kroeger B., Schroeder H., Zelder O., Haberhauer G., Kim J.
PI
Lee H., Hwang B;

DR WPI; 2001-582269/65.
DR P-PSDB; AAU71913.

Corynebacterium and Brevibacterium -
glutamicum, useful for producing methionine and lysine in
nucleic acids encoding metabolic pathway proteins from Corynebacterium

PS Disclosure; Page 281-283; 316pp; English.
XX

The present invention relates to the isolation of novel *Corynebacterium glutamicum* genes encoding metabolic pathway (MP) proteins (AAU71863-AAU71922). The metabolic pathway proteins of the invention include enzymes involved in the lysine and methionine biosynthetic pathway⁹. The polynucleotide sequences of the invention can be used for the large-scale production and/or modulation of expression of fine chemicals such as lysine and methionine. The sequences of the invention may be used to identify *C. glutamicum* and related organisms e.g. *C. diphtheriae* in a subject to detect diphtheria. AAS96073-AAAS96132 represent *C. glutamicum* genes encoding the novel metabolic pathway proteins of the invention.

sequence 1396 BP; 309 A; 396 C; 384 G; 307 T; 0 other;

Query Match	36.28;	Score 529.6;	DB 23;	Length 1396;
Best Local Similarity	64.33;			
Matches 828; Conservative				

32 tcgagtagaagaattgcgacatcaacctcattgttcgcgtaaggaacttaccctg 91

Db 114 tgaacttaaggttgcgcatcttcaactagcagagcagacgttcaaccagatctgttg 173
Qy 92 ctgagaaagaaatgcaggttctatgttctctgtgagcgttctccgttcaagccat 151
Db 174 cagagtatgagatgcaggttctatgttctctgtgagcgttctccgttcaagccat 233
Qy 152 tgaaggtgtcgaatcttctgttctccgttcaagcagcagcagcagcagcagcagc 211
Db 234 tgaaggtgtcgaatcttctgttctccgttcaagcagcagcagcagcagcagcagc 293
Qy 212 agacactcaagctctgtgtgtcagatgagcagcagcagcagcagcagcagcagcagc 271
Db 294 agacactcaagctctgtgtgtcagatgagcagcagcagcagcagcagcagcagcagc 353
Qy 272 aagatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 331
Db 354 agatagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 410
Qy 332 gataccagcttctgtgtgtcagatgagcagcagcagcagcagcagcagcagcagcagc 391
Db 411 gcttccagcttctgtgtgtcagatgagcagcagcagcagcagcagcagcagcagcagc 470
Qy 392 ggcctctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 451
Db 471 agatcttcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 527
Qy 452 ctacactctcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 511
Db 538 ccacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 587
Qy 512 cagaagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 571
Db 588 aggcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 647
Qy 572 aagacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 631
Db 648 cagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 707
Qy 632 caacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 691
Db 708 ccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 767
Qy 692 tcaacgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 751
Db 768 tgaacgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 827
Qy 752 cccttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 811
Db 828 cccttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 887
Qy 812 tcaatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 871
Db 888 tctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 947
Qy 872 ggcgttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 931
Db 948 ggcgttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1007
Qy 932 aggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 991
Db 1008 ctgtgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1067
Qy 992 actcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1051
Db 1068 acaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1127
Qy 1052 acatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1111
Db 1128 acatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1187
Qy 1112 agcac 1171
Db 1188 ccgcac 1247

Qy 1172 tcttctctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1231
Db 1248 tcttctctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1307
Qy 1232 tcatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1291
Db 1308 tcatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1367
Qy 1292 gaaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1319
Db 1368 gaaatctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1395

RESULT 10
AAC44037
ID AAC44037 standard; DNA: 1865 BP.
XX AAC44037;
XX 18-OCT-2000 (first entry)
DT
XX
DE Zea mays DNA fragment SEQ ID NO: 41385.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134216.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135621.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137752.
PR 07-JUN-1999; 99US-0137722.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139455.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142005.
PR 08-JUL-1999; 99US-0142305.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 26-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 34.8%; Score 508.8; DB 21; Length 1865;
Best Local Similarity 62.0%; Pred. No. 6.8e-140;

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151920.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155436.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.


```

PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 25-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 34.7%; Score 506.6; DB 21; Length 1727;
Best Local Similarity 61.9%; Pred. No. 2.9e-139;
Matches 897; Conservative 0; Mismatches 504; Indels 48; Gaps 4;

QY 34 gagaacgaatgacgacatcaactcactgctcgcgcgctgaaggaacttaccctgtg 93
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 106 gaatacaagatgaacgaatgctcaagcgaattcgcgtcgtcgaactcgaagctcgcgc 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 94 gagaaggaatgcagagctctatggtcttcgttgagcgttaltccgcttctaagcattg 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 166 gaagtgtgagatgcctggaactatgctgtctgacccaattcgaaccttctcagcattc 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 aaggggtgcagaatctctggtctcctccacatgacagctcagaacgcgtccctcattcag 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 aagagcgtatgaatcacggaatctctcactgacacccaacacgcgcgtacacacacga 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 214 aacactcaagctctggtctgctgacatgagctcctcctcgaacactctctcactacaa 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 266 aacctcaactgctcctcgtgctgacatgagctggtctcctcgaacactctctcactacaa 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 gatacagccgctgctgctatcgttctgcgcgcacacagcgaacacacagaagcagccgct 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 gaccacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 380
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 atccagcttctgcgcctggaagcgcgaacacacacacacacacacacacacacacacacac 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 381 --cgctgcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 394 gctctacatgagcagatgctcagagccacagcagctgtctcagatgctgtgtgtgtgtgt 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 439 gctctagatgtgggtccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 454 aacactcactatcccaaggctcga-----ttcgaacagccggtgtgtgtgtgtgtgtgt 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 499 actcttctgattcaatgaggtgtcttaagcgtgagagatctcttgaagaactgtgtcaagtt 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 ccagagccaacagaagctgacacactcgaatlacgcgtcgtcttctgtacacacacacag 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 559 cctgactactacttactactactactactactactactactactactactactactactact 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 gttctcaaccaagaacgaacacactgtgacacagtttctgcgcgcgtgaaggtgtgtgtgt 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 619 ggtctcaagttgtatccttaagaagttacacagaatgaagaggaactgtgtgtgtgtgtgt 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 622 gaagagacaacacaggtgtccacgcgcgtcactcagatcagagagagagagagagagagag 681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 679 gaagaaactcaactcgtgtgttaagagcttaccagatgcagcaaatggaactccttctgt 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 682 ttccacagcactcaacagcagcagcgtgttaacaagtccaagttccogataacatctacgcgc 741
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 739 ttcccttgcatttaagcttactgttccacagaagcagaacttgcgaacatttgaatgtgt 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 742 tgcgcacactcccttactcgttgtatcaacgcgttcttcgcgtgtcatgtcgcgcgcgcgc 801
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 799 tgcgttactcactccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 802 aacagctcgtcagtggttcaagcagatgtcgcgaagggctgcgtcactccctccgtgtgtgt 861
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 859 gttgctgttatctgtgtgatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 918
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 862 caagcgcgcgttatacatcacaagaactcgaacccaactgtcgtctcccaagctgcacatg 921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 919 gctgtgtcagagatcattgtgactgtgagattgatccactcgtgccttcaagcttctgtgt 978
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 922 gaaggtacacaggtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 981
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 979 gaaggaactcaagttcttacccttgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1038
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 982 tgcacgggaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1041
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1039 accacgcgttacaagaacacacacacacacacacacacacacacacacacacacacacac 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1042 attgtcgttacaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1099 attgtgtgcaacattgttcaacttctgacacatgtgacatgtgtgtgtgtgtgtgtgtgt 1158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1102 ccagcgcacacacacacacacacacacacacacacacacacacacacacacacacacac 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1159 cctgtgtgtgaagcgtatcacacacacacacacacacacacacacacacacacacacac 1218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1162 cagcgt---atcctcctctgtcgtgagggccgccttctaactgtgtgtgtgtgtgtgtgt 1218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1219 aaggtcgtgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1219 caaccatcctcgttattgtcaatgtcaatcaacacacacacacacacacacacacacacac 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1279 caccacagttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1279 t-----acgaaagagagagagacacacacacacacacacacacacacacacacacacac 1332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1339 tgaagacagagaagacagagagagagagagagagagagagagagagagagagagagagag 1398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1333 gatgaagaagtcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1399 gatgaagaagtcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1393 cagaagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1459 aaggaaccaatcgtactacacacacacacacacacacacacacacacacacacacacac 1518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1453 cgtataaa 1461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1519 aggtactga 1527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
AAC39836
ID AAC39836 standard; DNA; 1728 BP.
XX
XX AAC39836;
AC
AC 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26081.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
OS
OS EP1033405-A2.
PN
PN 06-SEP-2000.
PD
PD 25-FEB-2000; 2000EP-0301439.
PF
PF 25-FEB-1999; 9905-0121825.
XX
XX 05-MAR-1999; 9905-0123180.
PR
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.

```

PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130044.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 04-MAY-1999; 99US-0132048.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 07-MAY-1999; 99US-0132486.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 01-JUN-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0138899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142803.
 PR 12-JUL-1999; 99US-0142920.
 PR 13-JUL-1999; 99US-0142927.
 PR 14-JUL-1999; 99US-0143542.
 PR 15-JUL-1999; 99US-0143624.
 PR 16-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144532.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 16-AUG-1999; 99US-0148684.
 PR 17-AUG-1999; 99US-0149368.
 PR 18-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999: 99US-0158369.
 PR 13-OCT-1999: 99US-0159293.
 PR 13-OCT-1999: 99US-0159294.
 PR 13-OCT-1999: 99US-0159295.
 PR 14-OCT-1999: 99US-0159329.
 PR 14-OCT-1999: 99US-0159330.
 PR 14-OCT-1999: 99US-0159331.
 PR 14-OCT-1999: 99US-0159637.
 PR 14-OCT-1999: 99US-0159638.
 PR 18-OCT-1999: 99US-0159584.
 PR 21-OCT-1999: 99US-0160741.
 PR 21-OCT-1999: 99US-0160767.
 PR 21-OCT-1999: 99US-0160768.
 PR 21-OCT-1999: 99US-0160770.
 PR 21-OCT-1999: 99US-0160814.
 PR 21-OCT-1999: 99US-0160815.
 PR 22-OCT-1999: 99US-0160980.
 PR 22-OCT-1999: 99US-0160981.
 PR 22-OCT-1999: 99US-0160989.
 PR 25-OCT-1999: 99US-0161404.
 PR 25-OCT-1999: 99US-0161405.
 PR 25-OCT-1999: 99US-0161406.
 PR 26-OCT-1999: 99US-0161359.
 PR 26-OCT-1999: 99US-0161360.
 PR 26-OCT-1999: 99US-0161361.
 PR 28-OCT-1999: 99US-0161920.
 PR 28-OCT-1999: 99US-0161992.
 PR 28-OCT-1999: 99US-0161993.
 PR 29-OCT-1999: 99US-0162142.

Query Match 34.1%; Score 497.8; DB 21; Length 1728;
 Best Local Similarity 61.9%; Pred. No. 1.2e-136;
 Matches 898; Conservative 0; Mismatches 503; Indels 49; Gaps 5;

QY 34 gagtacaagattgcccagatcaactcattgtctgcgcgttaaggaacttacccttgc 93
 DB 104 gaatacaagttcaagaagacatgctctcaagccgattcgtctctgcgaactcgaagctgcgc 163
 QY 94 gagaagaagaatgcccaggtcttattgtcttcttgagcgttatctccgttactaagccattg 153
 DB 164 gaagttagatgctcggagatcagctgtctgtacacgaattcggactcttccagcatc 223
 QY 154 aaggtgtcagaatctctgtctcctccacacatgacagttacagacgcgtctccatcgag 213
 DB 224 aaagcgtcagatcaacacgcgactcttccacacacacacacacacacacacacacacacac 283
 QY 214 aactcaacagctcgttgcgtctgtatgacagtgagtggtctccgcaacatctcttccacaa 273
 DB 284 accctaactgctcgtcgtctgtatgacagtgagtggtctccgcaacacacacacacacacac 343
 QY 274 gatacagccgctgctgtcgtctgtctgtcgcgcacacagacacacacacacacacacacac 333
 DB 344 gacacagccgctcgtcgtctgtctgtcgcgcacacacacacacacacacacacacacacac 378
 QY 334 atcccaagcttgcgtcgtgaggaagcgaacacacacacacacacacacacacacacacac 393
 DB 379 --cgtgttgcgtcgtcgtgaggaagcgaacacacacacacacacacacacacacacacac 436
 QY 394 gcttcaacatgagcagatgctcaagagccacacagcagtgctgacagtgagtgagtgagtg 453
 DB 437 gctcctagatggtggtcagagtggtggtcgtcgtatgctgagtgagtgagtgagtgagtg 496
 QY 454 acaactcctacatcccaagggcttcga-----tcgaaacacgcggtgctgtt 501
 DB 497 actctttgtatcataggggtgttaagctgagagatcattgaaagacactgtaaggtt 556
 QY 502 ccagagccacaaagaagttgcaac 561
 DB 557 cctgattcctactctactgataac 616
 QY 562 gtctcaacacagacagac 621
 DB 617 ggtctcaagttgattcctatagaagtaacacacacacacacacacacacacacacacacacac 676

QY 622 gaagagacaacaacaggtgtccacacgcgtctcaacagctcgaagaaggaggaacacacac 681
 DB 677 gagaactaccac 736
 QY 682 ttccac 741
 DB 737 ttccctgcac 796
 QY 742 tgcgcac 801
 DB 797 tgcgcac 856
 QY 802 acaactcgtcattggttgcagcagatgctcgcgaagggcgtgcacacacacacacacacac 861
 DB 857 gttcgttattctgttgatagtgatggtgttggaaggggtgtgtgtgtgtgtgtgtgtgtgtgt 916
 QY 862 caagcgtcgtcgttattcacaac 921
 DB 917 gctgtgtcagatcattgtgtatgattgattcaccacacacacacacacacacacacacacac 976
 QY 922 gaagcctac 981
 DB 977 gaagcactcaggttcttccac 1036
 QY 982 tgcacagac 1041
 DB 1037 accac 1096
 QY 1042 attgtcgttacaac 1101
 DB 1097 attgtgtcacaac 1156
 QY 1102 ccagcatalcaac 1161
 DB 1157 cctgtgtgagatgattcaccac 1216
 QY 1162 caagct---atctccttctgtcgtgagggcgcgtcttcaacacacacacacacacacacac 1218
 DB 1217 aaggtcgtgac 1276
 QY 1219 caccac 1278
 DB 1277 caccac 1336
 QY 1279 t-----acgaagaagaggaatcctcgaagaaggtttacacacacacacacacacacac 1332
 DB 1337 tgaacagagaagaac 1396
 QY 1333 gatgaagaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1392
 DB 1397 gatgaagaaggttgcattactcacttgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1456
 QY 1393 cagaagcag-gctgactac 1451
 DB 1457 aaggaac 1516
 QY 1452 ccgttattaa 1461
 DB 1517 caagttactga 1526

RESULT 13

AAT44513
 ID AAT44513 standard; cDNA: 1767 BP.

AC AAT44513;

DT 22-FEB-1997 (first entry)

DE Atparagus S-adenosyl-L-homocysteine hydrolase cDNA clone.

XX S-adenosyl-L-homocysteine hydrolase; SHH; promoter; monocot; dicot;

transgenic plant; disease resistance; pathogen resistance; ds.
 XX Asparagus officinalis.
 OS
 XX Key Location/Qualifiers
 FH CDS 26..1483
 FT primer_bind /tag= a
 FT complement (281..303)
 FT primer_bind /tag= b
 FT /note= "PCR-1 primer"
 FT primer_bind 722..744
 FT /tag= c
 FT /note= "PCR-2 primer"
 XX
 PN W09632488-A1.
 PD 17-OCR-1996.
 XX
 PF 10-APR-1996; 96WO-GB00882.
 XX
 PR 10-APR-1995; 95GB-0007381.
 XX
 PA (ZENNE) ZENNECA LTD.
 PI Draper J, Greenland AJ, Skipsey M, Warner S;
 DR WPI: 1996-477138/47.
 DR P-PSDB; AAM01458.
 XX
 XX S-adenosyl-L-homocysteine hydrolase promoter - used for driving
 PT expression of effector genes, such as pathogen resistance genes, in
 XX transgenic plants
 XX
 PS Example 1; Fig 1; 57pp; English.
 XX
 CC A cDNA clone (AAT44513) codes for the S-adenosyl-L-homocysteine
 CC hydrolase (SHH) (AAM01458) of asparagus. A cDNA clone, designated
 CC DB6, was isolated from an asparagus cDNA library by homology
 CC searching, and the full-length sequence was obtained by library
 CC screening using DB6 as probe. The asparagus SHH sequence was used
 CC to obtain SHH genes from other plant species. Analysis of the
 CC Arabidopsis thaliana SHH gene (AAT44515) revealed a 1849 bp promoter
 CC (AAT44514) which can be used to drive expression of effector (e.g.
 CC pathogen resistance) genes in transgenic dicot and monocot plants.
 CC
 XX Sequence 1767 BP; 426 A; 462 C; 459 G; 420 T; 0 other:
 SQ
 Query Match 33.5%; Score 489; DB 17; Length 1767;
 Best Local Similarity 61.1%; Pred. No. 4.8e-134;
 Matches 886; Conservative 0; Mismatches 515; Indels 48; Gaps 4;

DB 337 -----ctccgtcttcgcgcgggaagggtcagaccctccagagatcgttggtgcaccgagcg 393
 QY 393 cgccttcacatggtccagatgtgtcagaagccacagaggttgctcagatgagtgatgc 452
 DB 394 tgccttcgactggggcccgcggtgtgacctcactcactcgtcgtatgacgagcgagac 453
 QY 453 taacctctatctccaaaggctt-----cgaaattcgaaacgcccgtgtcgt 500
 DB 454 cactctcttgatccatgaggggtgtgaagccgagagaagatcagagaagcgaggagat 513
 QY 501 tccagagcccaagagaagctcgaaccccgatccgacgtctcttgctacatcaagca 560
 DB 514 gcccgatccggcgtctcaccgaacatgcttgatccagatcgtgcacacatccagga 573
 QY 561 ggtcttaacccaagaacagacacactgtgacacagttgtcgcgcgacgtatgc 620
 DB 574 tgggtcctcaaggttgagcccccacagatcaggaagatgaaagtatgattgcgtgtc 633
 QY 621 cgaagagacaacacagaggtgtccacgcctcctaccagctcgaagaaggagcaactcct 680
 DB 634 ggaagagaccacacacggggttcagagagcttaccagatgacaggtacacatccctct 693
 QY 681 ctcccaagccatcaacgtcgaacagcgtgttacaaagctcgaagttcgaataatccag 740
 DB 694 ttccctcgatcaatgtcaatgtacgtcgcacacaaagacaagatttgacatctgtatg 753
 QY 741 ctgcgcacactccctatcgtatggtatcaacgcgtgttcgagttcagatcagtcgagca 800
 DB 754 atgcgcgacactcctcccgatgtgtcgtgagagggccacatgattgattgtgtcga 813
 QY 801 gacagctcgtcgtatggttgcagcgatgtgcgcaagggtcgtcgtcaatccctccgtg 860
 DB 814 ggttcagttgtcgtcggtgtatggtgtatgtcgcgagaggtcgtgtcgtcgaacagca 873
 QY 861 ccaagcgctcgtcgtatcatalcacaagaactccccaactcgtcgtccacgtcgtac 920
 DB 874 ggttggtcccggtatgtatgtgacgagatgcgcccactgtgtccttcaagccat 933
 QY 921 ggaaggtcaacaggttcgcgcgcgtatcgaagaatcgtcgaagaatcgtatcctgtac 980
 DB 934 ggaaggtcctcgcgtcctcaccctcgaagatgtgttcgcagagcgagatattctgtac 993
 QY 981 atgcacaggaactcgtatcactcgttgcagatgtgcccaggtgaaagataagc 1040
 DB 994 caccacgggttaacaagacatatactcgtcgcacacagatgaaagataagaacaatgc 1053
 QY 1041 tatgtcgttaacatcgcgcacatcgtatcgaatacgaatgtacagatgtcctatgaata 1100
 DB 1054 catgtctgcacaacttgcacttgcacacagagattgcacatgtagtlttgagata 1113
 QY 1101 cccgggtcacaagcacaatcccaatcaagcagaatcagacatgtgggaattcccaatg 1160
 DB 1114 cccgtgcacaaagaaacacacatccaaagcccaagctcgcgtgtggtcctcctaaac 1173
 QY 1161 ccac-----gcatcctcctctgtcgtgagggccgcttccaaactcgtcgtcgaag 1217
 DB 1174 caaacactggtataattgttctgtcgtgagggccgactcagtaacctgggtgtgcactg 1233
 QY 1218 tcaaccatcttctgttaagtataatgtcattcaaacacacgaacatcgtcgaactcgaact 1277
 DB 1234 tcaaccacagcttcttcatgtcctcgtcctcctcacaacaggtatgtcgaactgaagtt 1293
 QY 1278 ctacgaaag-----agaggaatctcgaagaaggttcaacaactcccgagatct 1331
 DB 1294 gtggaatgagaagcaagcgcgcaagttatgagaagaagttatcgtcctcccaagcatct 1353
 QY 1332 cgaatgaagaggtcgtcgtccctcgaactcgtcgaatcgtatgcacactcaagaactac 1391
 DB 1354 tgaatgaaggttgaacgctcctcactgtggaaggtcgaagaagcttaacaaggtcag 1413
 QY 1392 acagaagcaggtgactacatcaacggttccagttgaggtgtccttaagatcgtatgctta 1451

Db 1414 cccttcacagcgactacatcagctcccatcgaggtccctcctacaagccactca 1473
 QY 1452 cccttattat 1460
 Db 1474 caaggtacta 1482

RESULT 14
 AAQ29419
 ID AAQ29419 standard; mRNA: 1812 BP.
 AC AAQ29419;
 XX 10-MAR-1993 (first entry)
 DE Nicotiana tabacum gene expressing at floral differentiation.
 XX
 KW Flower; induction; ss.
 XX
 OS Nicotiana tabacum.
 XX
 XX Key Location/Qualifiers
 FH CDS 61..1521
 FT /*tag= a
 XX
 PN JP04258292-A.
 XX
 PD 14-SEP-1992.
 XX
 PE 14-FEB-1991; 91JP-0020702.
 XX
 PR 14-FEB-1991; 91JP-0020702.
 XX
 PA (NISB) JAPAN TOBACCO INC.
 DR WPI: 1992-354683/43.
 XX P-PSDB: AAR26500.
 PT Gene expressing at floral differentiation for flowering control -
 PT Oobd. from culture of floral axis epithelium cells of Nicotiana
 PT tabacum, for introduction into other plants
 XX
 PS Claim 1; Page 7; 8pp; Japanese.
 CC Epithelial sectional of floral axis of Nicotiana tabacum BY-4 were
 CC cultured in Murashige-Skoog medium contg. kinetin and indoleacetic
 CC acid. mRNA was extracted from the cultured cells by conventional
 CC methods and a cDNA library prepd. mRNA extracted from untreated
 CC floral axis and the above mRNA were used to prep. corresp. cDNA
 CC probes which were used to screen the cDNA library. The clone from
 CC untreated floral axis mRNA did not hybridise but the treated mRNA
 CC probe hybridised to a clone of 1,812 bp. The clone expressed at
 CC floral differentiation. The gene can be introduced into other
 CC plants or can be suppressed by an antisense technique for the control
 CC of flowering of plants.
 CC
 XX
 XX Sequence 1812 BP; 468 A; 377 C; 452 G; 515 T; 0 other:
 SO

Query Match 33.4%; Score 488.4; DB 13; Length 1812;
 Best Local Similarity 61.18; Pred. No. 7.3e-134;
 Matches 886; Conservative 0; Mismatches 516; Indels 48; Gaps 4;

QY 33 cgaatcacagaattcgacatcaacctccatgcttcctcgccgtaaggaactaaccttcg 92
 Db 99 cgaatcacagaattcgacatcaacctccatgcttcctcgccgtaaggaactaaccttcg 158
 QY 93 tgaagaagaatcgacatcaacctccatgcttcctcgccgtaaggaactaaccttcg 152
 Db 159 cgaatcgacatcaacctccatgcttcctcgccgtaaggaactaaccttcg 218
 QY 153 gaaggggtcgacatcaacctccatgcttcctcgccgtaaggaactaaccttcg 212

Db 219 taaggtgtaagattactcgtgatcttatacatatgaccattcaactgcaatttgatgta 278
 QY 213 gacactcaagctcttggtggtggtatgacatgaggtcttcctcgcaactctctctacaca 272
 Db 279 aaccttactgcttggtggtggtatgacatgaggtcttcctcgcaactctctctacaca 338
 QY 273 agatacagcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 332
 Db 339 agatacagcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 371

QY 333 tatccagctcttcgctgtaagggcggaacactccgaatctggtggaacacataccg 392
 Db 372 cgcgcgctgctgctgtaagggcggaacactccgaatctggtggaacacataccg 431
 QY 393 cgcctcacaatcgcaagatgtaagggcggaacactccgaatctggtggaacacataccg 452
 Db 432 ggcacttgactggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 491

QY 453 tacactcctcatctccaaaggtctggaatccgaacagcgctgctgctgctgctgctgctg 500
 Db 492 tacactcctcatctccaaaggtctggaatccgaacagcgctgctgctgctgctgctgctg 551

QY 501 tcaagggccaacaggaactgacacacactccgaatctggtggaacacataccg 560
 Db 552 cccaagctcctaactctacacacacactccgaatctggtggaacacataccg 611

QY 561 ggtcttcaacacaggaacacacacactccgaatctggtggaacacataccg 620
 Db 612 aagtttggaagacgactcctttaaataatccaaagatggaagaaactcgtcggtgcttc 671

QY 621 cgaaggaacacacacacacactccgaatctggtggaacacataccg 680
 Db 672 tgaaggaacacacacacacactccgaatctggtggaacacataccg 731

QY 681 ctccccaacacacacacacacactccgaatctggtggaacacataccg 740
 Db 732 ttccctgcatatattgtatgtcttctgtaacaaagcaagcttcgaacactgtaacg 791

QY 741 ctgcgcgaactcccttatcgtatgatacaacgctgcttcggtgctgctgctgctgctgctg 800
 Db 792 atgcgcgaactcccttatcgtatgatacaacgctgcttcggtgctgctgctgctgctgctg 851

QY 801 gacagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 860
 Db 852 ggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 911

QY 861 ccaagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 920
 Db 912 agccggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 971

QY 921 ggaagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 980
 Db 972 ggaagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1031

QY 981 atgacaggaactgcatatcatctctgttgacatgagtggccagatgaagataagcg 1040
 Db 1032 caagcaaggtatcaaggaacattatcatggttgacacatgagaagatgaagaacatgctg 1091

QY 1041 tatgtcggtatacatcgccacttgcatacgaatattgatacagatggcctcatgaata 1100
 Db 1092 catgttgcaacatttgctacatttgcaacagaaatcgacatgctggttcgagaccta 1151

QY 1101 cccaagggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1160
 Db 1152 cccgggtgctcaaggaatgacacatttaagcctcaacacagaaagatggcttcctcgacac 1211

QY 1161 ccaac---gtatcctcctctctgctgctgagggcgctcttcaactctgctgctgctgctg 1217
 Db 1212 caacaggtgcatcatgtcttctgctgagggctgcttcatatgaacttggatggtgacag 1271

QY 1218 tcaccacatcttcgttatgtaatgcatcacaacacagacactcgtcagctcagact 1277
 Db 1272 acaccctagtttgatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1331

QY 1278 ct-----acgaagaagaggaatctcgagagaaggtttacacacttcggaagcact 1331
 Db 1332 gtggaatggaagaagcagtggaagatggaagaagatgtcttgcacccacact 1391
 QY 1332 cgaatgaagaatgcctgcctccacactgcgtatcgtatgcacacttacaagcttac 1391
 Db 1392 cgaagagaaggttgcctgcctccacactgcgtatcgtatgcacacttacaagcttac 1451
 QY 1392 acagaagcaggtgactacatcaacgttcacagttgaggttccttaacagctgactta 1451
 Db 1452 gaaggatcaagctgactacatcaacgttcacagttgaggttccttaacagctgactta 1511
 QY 1452 ccgttatata 1461
 Db 1512 caggtactga 1521

RESULT 15

AAT29194
 ID AAT29194 standard; cDNA to mRNA; 1812 BP.

AC AAT29194;

DT 11-NOV-1996 (first entry)

DE S-adenosylhomocysteine hydrolase gene.

KW S-adenosylhomocysteine hydrolase; transgenic plant; growth;
 resistance; antisense; ribozyme; protein binding; tobacco; tomato;
 rice; maize; ss.

OS Nicotiana tabacum.

Key Location/Qualifiers
 CDS 64..1521

FT /product= S-adenosylhomocysteine hydrolase.
 FT /note= "Although not indicated in the text of the
 FT specification, the indexer believes this to
 FT be the most likely position for the CDS."

PN W09614734-A1.

PD 23-MAY-1996.

PE 15-NOV-1995; 95WO-JP02333.

PR 15-NOV-1994; 94JP-0304200.

PA (NISR) JAPAN TOBACCO INC.

PI Kuwata S, Masuta C, Tanaka H, Uehara K;

DR WPI: 1996-259465/26.

PT Transformed plants with inhibited expression of SAHH gene - have
 PT improved growth characteristics and viral resistance

PS Disclosure: Page 32-33; 48pp; Japanese.

CC Transgenic plants having substantially inhibited expression of the
 CC genomic S-adenosylhomocysteine hydrolase (SAHH) gene have improved
 CC growth characteristics and viral resistance. The plants, such as
 CC tobacco, tomato, rice and maize, inhibit the SAHH gene by the
 CC production of antisense RNA, by containing the SAHH gene fragments
 CC of the SAHH gene and/or by containing a nucleotide sequence encoding
 CC a protein which binds to SAHH.

XX Sequence 1812 BP; 471 A; 374 C; 453 G; 514 T; 0 other;

Query Match

33.4%; Score 486.4; DB 17; Length 1812;

Best Local Similarity 61.1%; Pred. No. 7 3e-134;
 Matches 886; Conservative 0; Mismatches 516; Indels 48; Gaps 4;

QY 33 cgaatgaagaatgacgaatcaactcactatgtctgcgcgaagaactaccctgctg 92
 Db 99 cgaatgaagaatgacgaatcaactcactatgtctgcgcgaagaactaccctgctg 158
 QY 93 tgaagaagaatgacgaatcaactcactatgtctgcgcgaagaactaccctgctg 152
 Db 159 cgaatgaagaatgacgaatcaactcactatgtctgcgcgaagaactaccctgctg 218
 QY 153 gaaggtgtcagaatcctcctcctcctcctcctcctcctcctcctcctcctcctc 212
 Db 219 taaagtgtaagatcactgactcctcctcctcctcctcctcctcctcctcctcct 278
 QY 213 gacactcactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 272
 Db 279 aacctcactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 338
 QY 273 agatacagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 332
 Db 339 agatcagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 371
 QY 333 tatccagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 392
 Db 372 cgcgcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 431
 QY 393 cgcctcaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 452
 Db 432 ggcctcaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 491
 QY 453 tacactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 500
 Db 492 tacactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 551
 QY 501 tccagagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 560
 Db 552 cccagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 611
 QY 561 ggtctcaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 620
 Db 612 aagttgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 671
 QY 621 cgaagaagaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 680
 Db 672 tgaagaagaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 731
 QY 681 ctccagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 740
 Db 732 ttcctcgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 791
 QY 741 ctgcgcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 800
 Db 792 atgcgcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 851
 QY 801 gaaagctcgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 860
 Db 852 ggttcgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcga 911
 QY 861 ccaagcgtcgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 920
 Db 912 agccgtggtcgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 971
 QY 921 ggaagtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 980
 Db 972 ggaagtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 1031
 QY 981 atgcagaagaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 1040
 Db 1032 cagcagaagaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 1091
 QY 1041 tatgtcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgagcgaatgag 1100

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2002, 01:48:22 ; Search time 2720.84 Seconds

(without alignments)
11236.850 Million cell updates/sec

Title: US-09-759-990-1

Perfect score: 1461

Sequence: 1 atggctgcgaatacactac.....ctgatgcttaccgtattaa 1461

Scoring table: IDENTITY_NUC

Gap 10.0 , Gape 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_da:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pal:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_da:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hlg_hum:*
- 31: em_hlg_inv:*
- 32: em_hlg_other:*
- 33: em_hlg_inv:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

1	1459.4	99.9	1461	6	AX195549	AX195549 Sequence
2	1440.2	98.6	1882	3	TYV40872	U40872 Trichomonas
3	655.2	44.8	37898	1	SCB34	AL35862 Streptomy
4	615.2	42.1	3800	1	SAR416377	AJ416377 Streptomy
5	596.4	40.8	1557	6	AX063937	AX063937 Sequence
6	596.4	40.8	1557	6	AX244105	AX244105 Sequence
7	596.4	40.8	1557	6	AX127145	AX127145 Sequence
8	594	40.7	1422	6	AX120920	AX120920 Sequence
9	535.4	36.6	1488	1	AF262755	AF262755 Mycobacte
10	535.4	36.6	1488	1	AX023852	AX023852 Sequence
11	535.4	36.6	18249	6	AE007145	AE007145 Mycobacte
12	535.4	36.6	18249	6	AE007145	AE007145 Mycobacte
13	529.6	36.2	1396	6	AX063941	AX063941 Mycobacteri
14	529.6	36.2	1396	6	AX244109	AX244109 Sequence
15	529.6	36.2	1396	6	AX244109	AX244109 Sequence
16	512.6	35.1	1837	8	WHTSHR	LI1872 Trifolium ae
17	511.4	35.0	312050	1	MCU79766	U79766 Mesembryant
18	506.6	34.7	1729	8	MLEPRN3	AL583919 Mycobacte
19	506.6	34.7	1802	8	AY042866	AY042866 Arabidops
20	506.6	34.7	1802	8	AF059581	AF059581 Arabidops
21	506.6	34.7	1821	8	AF325037	AF325037 Arabidops
22	506	34.6	13407	1	AE008877	AE008877 Arabidops
23	506	34.6	13407	1	AE007946	AE007946 Agrobacte
24	501.8	34.3	1798	8	PSSADHY	AE007946 Agrobacte
25	497	34.0	1790	8	ATSADLH	X79905 Phalaenopsi
26	490	33.5	1812	6	AR071283	Z97059 Arabidopsis
27	490	33.5	1812	6	RO3902	EO3902 A gene invo
28	490	33.5	1812	6	RO3902	D45204 Nicotiana t
29	489	33.5	1767	8	TOBSALHHH	AE0302 A gene invo
30	488.2	33.4	215050	1	AS7643	AE0302 A gene invo
31	483.6	33.1	2110	8	AL646057	AE0302 A gene invo
32	481.6	33.0	1791	8	TOBCHP57A	AE0302 A gene invo
33	475.6	32.6	10791	1	PUMSHHB	AE0302 A gene invo
34	473	32.4	1685	8	AE005699	AE005699 Caulobact
35	472.4	32.3	1754	8	AF428329	AF428329 Arabidops
36	471.4	32.3	1738	8	AF161705	AF161705 Lycopersi
37	469.8	32.2	1738	8	AY059888	AY059888 Arabidops
38	466.6	31.9	1697	8	AY050783	AY050783 Arabidops
39	465.6	31.9	1697	8	CRSAH1	Z26881 C. roseus SA
40	463.4	31.7	1824	8	TOBCHP57B	AF185635 Lupinus l
41	462.8	31.7	1824	8	AF185635	AF185635 Lupinus l
42	462.8	31.7	332635	1	ALFMSA25	L36119 Medicago sa
43	459.8	31.5	11711	1	AE003005	AE003005 Mesorhizo
44	459	31.4	1660	8	AE003941	AE003941 Xylella f
45	455.8	31.2	10481	1	AE009636	AF462828 Arabidops

ALIGNMENTS

RESULT 1

AX195549 1461 bp DNA linear PAT 28-AUG-2001

LOCUS AX195549

DEFINITION Sequence 1 from Patent WO0151651.

ACCESSION AX195549

VERSION AX195549.1 GI:15386074

KEYWORDS

SOURCE

ORGANISM

Trichomonas vaginalis.

Trichomonas vaginalis

Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;

Trichomonas.

1 (bases 1 to 1461)

Hoffman, R.M., Xu, M., and Han, Q.

High expression and production of high-specific activity

recombinant s-adenosyl homocysteine (sahh) and improved assays

for s-adenosylmethionine (sam)

Patent: WO 0151651-A 1 19-JUL-2001;

Anticancer, Inc. (US)

Location/Qualifiers

1..1461

FEATURES

source

BASE COUNT 372 a 427 c 335 g 327 t

ORIGIN

Query Match 99.9%; Score 1459.4; DB 6; Length 1461;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagctgcaatcactactggtggtcattggtacagattggccagatcaacc 60
 DB 1 ATGGCTTGCAAAATCAGCTACTGCTGCTCATTGAGTACAGATTGCCATCAACCTC 60
 QY 61 catgttcgagccgtaaggaactaaccttgcctgagaaagaaatgccaggtctatgatt 120
 DB 61 CATGTTCTCGGCCGTAAGGAATTACCTTGCTGAGAAAGAAATGCCAGAGCTTATAGTT 120
 QY 121 ctctgtagaggtatctccgtcttaagccaattgaaggtgtgaagatctctgttccctc 180
 DB 121 CTTCGTGAGGCTTATCTCGCTTCTAAGCCATTGAAGGGGTGTGAATCTCTGGTCCCTC 180
 QY 181 caatgaacatccagagacgagccgtctcactcagacactcacagctcttggtgtatg 240
 DB 181 CACATGACAGTCCAGACAGCCGCTCTCATCGACACTCACAGCTTGTGCTGATGTC 240
 QY 241 agatggagcttctcgaacatcttctcacaagaatcacagccgtgctgctatcgtgtc 300
 DB 241 AGATGGGCTTCTGCAACATCTTCTACACAAAGATACAGCCGCTGCTGATCGTTGTC 300
 QY 301 ggcaccaagagcaacccagagagagccagctatccagcttgcctgagaaaggagaa 360
 DB 301 GGCCCAACAGGACACCCAGAGAGCCAGCGGTATCCAGTCTTCGCTGGAAGGGCGAA 360
 QY 361 acactccagaatacttggagagacacataccgagctcactacatgagcagaaatg 420
 DB 361 ACACTCCAGAAATCTGGAGAGAACATATCCCGCTCTCACATGGGCGAGATGGCAAGGC 420
 QY 421 ccacagagagtttgcagatgagtggtgtgctacactcctcattccaaagggttga 480
 DB 421 CCACAGCAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 481 ttggaagagagcgt 540
 DB 481 TTGGAAGAGCGCGT 540
 QY 541 gtcttgcatactcaagcaggttctcaacccaagaagaacacatgagcagagttgct 600
 DB 541 GTTCTGTGCTACACTCAAGAGGCTTCAACCAAGAGAACACATGGGCGAGATGGCT 600
 QY 601 gccggatgaagcgt 660
 DB 601 GCCGGATGAAGCGT 660
 QY 661 gagaagagaggaacactccttccagcagcatcaacgtcaacagcagctgttacaagtcc 720
 DB 661 GAGAAAGAGGGCAAACTCTCTTCCAGCATCAACGTCAACAGCGCTGTACAAAGTCC 720
 QY 721 aagttgataactaactaagcgtgcgcacatcccttatcgatggtatcaacagcgtctcc 780
 DB 721 AAGTTGATTAACATCTACCGCTGCCGCCATCTCTTATCATGATGATTAACCGGCTTCC 780
 QY 781 gatgcatgatcggcgagcaagacgtctcgtcatggtgtgaagcgtgtgtgtgtgtgtgt 840
 DB 781 GATGTCATGATCGCGGAGAGACAGCTCTCTCATGGGTATACGGCGATGTGCGCAAGGCG 840
 QY 841 tgcgtcaatccctcgttgcagagcagcgtcgttatacatcacaagaactcgaccacatc 900
 DB 841 TGCCTCATATCCCTCGTGGCCAAAGCGCTCGCTTATCATACAGAAAGTGCACCAATC 900
 QY 901 tgcgtctccagcgtgcagatgaagcgtacagagtcgcgcgcatcggagagtgatcgaag 960
 DB 901 TGCCTCTCCAGCGTGCATGGAAGGCTACAGAGTCCGCCGATCGAGGAAGTGTGTCAAG 960
 QY 961 gatgcgatatctggttacaatcacaggaactgagatatactctgttgaatgag 1020
 DB 961 GATGCGATATCTGTTACATGACAGAGAACTGGAGATATCTCTGTTGACATGATG 1020

DB 961 GATGCGATATCTGTTACATGACAGAGAACTGGAGATATCTCTGTTGACATGATG 1020
 QY 1021 gccagatgaagagatgaagcgtatctgtgttaacatcgcgcacttcgataaagaattgat 1080
 DB 1021 GCCAGATGAAGAGATGAAGCTATTTGTCGTAATCGGCCACTTCCATTAAGAAATGAT 1080
 QY 1081 acaagagcctcatgaatcaccagagcatcaagcacatcccaatcaagcagaatcagac 1140
 DB 1081 ACAGATGCGCTCATGAATAATACCAGAGCATCAAGCATATCCCAATCAAGCCAGATACGAC 1140
 QY 1141 atgtaggaattccagatgagcagcgtatccctctctgtgtgagggcgccctctaac 1200
 DB 1141 ATGAGGAATTCCTAGATGGCCAGCTATCTCTCTTCTTCTAGAGGCCGCGCTTAAAC 1200
 QY 1201 ctgagctgcgtacagagtcaccatcttcgtatgltcaatgcatcacaanaacagaca 1260
 DB 1201 CTGAGCTGCGCTACAGATGATCCCATCTTCTGTTATGATGATGATGATGATGATGATGAT 1260
 QY 1261 ctgcgtcagctcgaactcagaaagagagaaatctcgagaaagaggttatacactt 1320
 DB 1261 CTGCTCAGCTCAGCTTACGAAAGAGAGAAATCTCGAAGAAAGGTTTACACACTT 1320
 QY 1321 ccgaagcattcgtatgaagaagtgcgtgcgtcgcctcagctcgatctcgtatgctcact 1380
 DB 1321 CCGAAGCATCTGATGAAGAAAGTGCCTGCTGCTCAGCTCGATCTCGATGTCACACTT 1380
 QY 1381 acaagcttacacagaagcagcgtgactacatacaagcttcaggttgaaggtctcacaag 1440
 DB 1381 ACAAGCTTACACAGAAAGAGGCTGACTATCATCAAGCTTCCAGTTGAGAGGCTTACAAAG 1440
 QY 1441 tctgattgttaccgttattaa 1461
 DB 1441 TCTGATGCTTACCGTTATTAA 1461

RESULT 2
 TVU40872 1882 bp DNA linear INV 31-OCT-1996
 LOCUS Trichomonas vaginalis S-adenosyl-L-homocysteine hydrolase gene,
 DEFINITION complete cds.
 ACCESSION U40872
 VERSION U40872.1 GI:1118008
 KEYWORDS Trichomonas vaginalis.
 SOURCE Trichomonas vaginalis.
 ORGANISM Eukaryota; Parabasalidea; Trichomonadida; Trichomonadidae;
 Trichomonas.
 REFERENCE 1 (bases 1 to 1882)
 AUTHORS Bagnara,A.S., Tucker,V.E., Minotto,L., Howes,E.R., Ko,G.A.,
 Edwards,M.R. and Daves,I.W.
 TITLE Molecular characterisation of adenosylhomocysteine from
 Trichomonas vaginalis
 JOURNAL Mol. Biochem. Parasitol. 81 (1), 1-11 (1996)
 MEDLINE 97047381
 REFERENCE 2 (bases 1 to 1882)
 AUTHORS Bagnara,A.S.
 TITLE Direct Submission
 JOURNAL Submitted (19-NOV-1995) Aldo S. Bagnara, Biochemistry and Molecular
 Genetics, The University of New South Wales, Sydney, New South
 Wales, 2052, Australia
 FEATURES
 source location/Qualifiers
 1..1882
 /organism="Trichomonas vaginalis"
 /strain="WAA38"
 /specific_host="Homo sapiens"
 /db_xref="taxon:5722"
 1..290
 291..1751
 /c_number="3.3.1.1"
 /codon_start=1
 /product="S-adenosyl-L-homocysteine hydrolase"
 /protein_id="AAC47319.1"
 /db_xref="GI:1118009"

/translation-"NACKSPAGAPPEYRIADILUHLVGRKELTAEKEMPGLVLRER
 TSASPLKGVRLSSLSLHMTVQTALETLTALGADVMACNIFSTODTAAALVVG
 TGPBPAIPEVFAWKGETLPEWENTYVLTALTPDSCGPOOVVDDGADATLLISKFE
 FETAGAVEPEPTADNLEYRCVLATLEKOVNODNHHMTVAGNVSEPTTGVHRLY
 OLEKGLLEPAIINVDATVSKSPDNVIGCRSLIDINGINASVNGIGKRALVNGYD
 VSGKAOOSLRGGARVITTEPNPICLQVMESTOYRRIEYKADVFETCTNCDI
 ISVDMQMDKRAITGNIGHTNEIDTGLMKPGIKHPDKEDYDMEPPDGAHLL
 LAEGRLNMGCTGHPSEFVMSFTNOLADLYERKGMLEMKVYVLPRLHDEEVVR
 LHLGSLDHLTKOKROADYINVPVEGPKSDAYRY"

3'UTR
 polyA-signal
 BASE COUNT 506 a 499 c 395 g 482 t
 ORIGIN

Query Match 98.6%; Score 1440.2; DB 3; Length 1882;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1448; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 atgcttgcgaatacactactgtgtcctcattcgaatgcgaatacactc 60
 Db 291 ATGGCTTCAATACACCTGCTGCTCATTGAGTACGAAATGCGCATCAACCTC 350
 QY 61 catgttcgcgcgaagaacttacccttgcctgagaagaattgcgaattcattgtt 120
 Db 351 CATGTTCGCGCCGTAAGAACTTACCTTGTCTGAGAAAGAAATGCCAGGCTTATG 410
 QY 121 ctctgtgagcgttattcgcctcctaagccattgaaggtgtcgaatctcgttccctc 180
 Db 411 CTCTGTGAGCGTTATTCCGCTCTTAAGCATGAAAGGTGTGAGAAATCTCGTTCCTC 470
 QY 181 cacatgacatccagaagacgcctcctcattcgaacactcgaactcgttgcctgtc 240
 Db 471 CACATGACATGACACAGCGGCTCTTATTTAGACACTCACACTCTTGTGCTGAGTGC 530
 QY 241 agatggccttcctcgaacatctctcctacagaatacagcgcctgtctatcgttgc 300
 Db 531 AGATGGGCTTCTGCAACATCTTCTTACAAAGATACAGCCGCTGCTATCTGTTGTC 590
 QY 301 ggcgcgaagcgaacacgaagaagcgaagcgttaccagcttcgtcgttgaaagcgaa 360
 Db 591 GGCCCAACAGGACACACGAGAGAGCCAGCGGTATCCAGTCTTCCGCTGGAAGGGGAA 650
 QY 361 acaatccagaatactggaagaacacatccgcgtcctcattcgaagcgttcgaagc 420
 Db 651 ACACTCCAGAACTTGGGAGACACATACCGGCTCTCAATGGCCAGATGGTCAAGGC 710
 QY 421 ccaagaagatgttcgaatgagtgatgtatcactcctcctcgaaggcctcga 480
 Db 711 CCACAGCAGGTTGTGATGATGATGTGTATGCTTACACTCTCTCAAGGCTTCGAA 770
 QY 481 ttggaacacgcgcgtgtcgttcgaagaacgaagaagctgaacactcgaatacgcgtgc 540
 Db 771 TTGGAACACGCGGTGTGCTCCAGAGCCACAGAGCTGACAACTCGAATACCGCTGC 830
 QY 541 gtcttctcactcaagcagcttcaacgaagaacgaagaacgcgtggaacaaagtgtc 600
 Db 831 GTTCTTCTACACTCAACGACAGGCTTCAACCAAGACAAACACACTGGGACACAGTTGCT 890
 QY 601 gccgcgaatgaagcgttcttcgaagaagaacaaacaggtgtccacgcctcctcagctc 660
 Db 891 GCCGCGATGAAAGGCTGTTTCGAAAGACAAACACAGGTCACACGCGCTCTACCAAGCTC 950
 QY 661 gagaagaagggaacacactcttccagccatcaacgcgtcaagaagcgtttacaagtc 720
 Db 951 GAAAGAGAGGCAAACTCTCTTCCACCATCAACAGCTCAACGACGCTTCAAAAGTCC 1010
 QY 721 aagtcgaatacactcagcgttcgcgcacccctcctcagatgattcaacagctgttc 780
 Db 1011 AAGTTGATTAACATCTACGCGCTGTGCGCACCTCTTATCGATGATCAACCGTCTTCC 1070
 QY 781 gatgtcatgattcgcgcgaagaagcctcgtcattggtttacgcgagatgtcgcgaagc 840

Db 1071 GATGTCATGATCGGCGCAACAGACGCTCTGTCATGGGTTACGGCGATGTCGGGAAGGC 1130
 QY 841 tgggttaattccctccgttgcgaagcgctgcgttattcattcagaactcgaaccatc 900
 Db 1131 TGGGCTCAATCCCTCCGCGCAAGCGCCTGCTTATCATCAAGAACTCAACCTTATC 1190
 QY 901 tggccttcgaagctgcattgaaagcgttaccaggtccgcgaatcgaagaagtcgaag 960
 Db 1191 TGGGCTCTCAAGGCTGTCATGGAAGGCTTACAGGTCGCGCGATCGAGGAAGTCTCAAG 1250
 QY 961 gatgtcatatcttcgttacaatgcagaagaactgcgaatcattcctcgttgcattgac 1020
 Db 1251 GATGTCATGATCTGTTACATGCAAGGAAGCTCGATATATCTCTTACATGATG 1310
 QY 1021 gcccaatgagaataagcgtattgtgtgtacatcgcgcacacttcgataagaattgat 1080
 Db 1311 GCCCAGATGAAGGATTAAGGCTATTGTGCTGAATCAATCGGCACTTCGATCAACAAATG 1370
 QY 1081 acagatggcctcattgaataaccaggaacacgaacacacacacacacacacacacac 1140
 Db 1371 ACAGATGGCCTCATGAATATACCGGATCAAGCATATCCCAATCAAGCAGATATGCGAC 1430
 QY 1141 atgtggaaatcccaagatgcgaacgtatcctcctccttgcgtgagggcgcctctaac 1200
 Db 1431 ATGTGGAAATTCACGATGCGCACGCTATCTCTTGTGAGGGCGGCTCTTAC 1490
 QY 1201 ctgtgctgcctcagatgcac 1260
 Db 1491 CTGTGTTGGCTTACAGGTACCCATCTTCTTGTATGTAATGTCATCAACAAACAGAA 1550
 QY 1261 ctgcctcagctgcacactcctcgaagaagaagaatcctcgaagaagaagtttaccact 1320
 Db 1551 CTGCTCAGGCTCAGCTCTACGAAAGAGAGAAATCTCGAATGAAGGTTTACACACTT 1610
 QY 1321 cogaagcatctcgaatgaagaagtcgtcgtcgtcctcactcgaatcctcgtatgcacact 1380
 Db 1611 CGAAGCATCTCGATGAAGAGGTTCTGCGCTCACACCTCGGATCTCTGATGTCCACTT 1670
 QY 1381 acaaaacttacacagaagcagcgtgactacatacactcgaatgaagaagttcctaac 1440
 Db 1671 ACAAACTTACACAGAAAGCGGCTGATCACTCAACGTTCCAGTGAAGGTCCTTCAAG 1730
 QY 1441 tctgatgttaccgttattaa 1461
 Db 1731 TCTGATGCTTACCGTTATTAA 1751

RESULT 3
 SCE34/c
 LOCUS 3789 bp DNA linear BCT 01-JUN-2000
 DEFINITION Streptomyces coelicolor cosmid E34.
 ACCESSION AL353862
 VERSION AL353862.1 GI:7649483
 KEYWORDS adenosylhomocysteinease; integral membrane protein; L-lactate permease; mannose-6-phosphatase isomerase; nucleoside phosphorylase; oxidoreductase; phosphomannomutase; rmd; secreted protein; transcriptional regulator; transport protein; whb.
 Streptomyces coelicolor A3(2).
 SOURCE
 ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomyces; Streptomyces; Streptomyces.
 1 (bases 1 to 3789)
 Redenbach, M., Kleser, H.M., Denapate, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A.
 A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome
 Mol. Microbiol. 21 (1), 77-96 (1996)
 JOURNAL MEDLINE 97000351
 2 (bases 1 to 3789)
 REFERENCES Saunders, D.C. and Harris, D.
 JOURNAL Unpublished
 3 (bases 1 to 3789)
 Cerdano, A.M., Parkhill, J., Barrell, B.G. and Randal, M.A.

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (25-APR-2000) Streptomyces coelicolor sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof. David A. Hopwood, [3] John Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk NR4 7UH, UK

Notes:
Streptomyces coelicolor sequencing at the Sanger Centre is funded by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.
(URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>)
CDS are numbered using the following system eg SC7B7.01c, SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand)
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon.
Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/>
Jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
IMPORANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.
Cosmid E34 Overlaps cosmid E33 on the AseI-E genomic restriction fragment.

FEATURES

source

misc_feature

rRNA

rRNA

rRNA

gene

CDS

1. 37898
/organism="Streptomyces coelicolor A3(2)"
/strain="A3(2)"
/db_xref="taxon:100226"
/clone="cosmid E34"
1. 161
/note="nominal overlap with Streptomyces coelicolor cosmid STE33"
complement(1..105)
/note="probable 23S rRNA (fragment), coordinates taken from similarity to SW:SARRND (EMBL:M27245) Streptomyces ambifaciens rrrd gene cluster encoding 16S, 23S and 5S ribosomal RNAs"
complement(1..2116)
/note="coordinates taken from similarity to SW:SARRND (EMBL:M27245) Streptomyces ambifaciens rrrd gene cluster encoding 16S, 23S and 5S ribosomal RNAs"
complement(2470..3477)
/gene="SCE34.01c"
complement(2470..3477)
/gene="SCE34.01c"
/note="SCE34.01c, possible integral membrane protein, len: 335 aa; similar to TR:069662 (EMBL:AL022121) Mycobacterium tuberculosis hypothetical 35.5 kD protein MTW025.042c, 330 aa; fasta scores: opt: 1119 z-score: 1285.3 E(): 0; 51.8% identity in 334 aa overlap. Contains possible hydrophobic membrane spanning regions"

RBS
RBS
CDSgene
CDSgene
CDS

/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB88904.1"
/db_xref="GI:7649484"
/translation="MDLDVFSAHRAEMDRDLARRRRITGPEDVLAALQRTAT
HLSLIQSSAPDQTLGRSLQVARRASVYGTFRASRVDRTRLTQGFPAVYRARHM
WPTALSTAAVALHGMWIGTHPEVOSTIAAPSELRELTRPGQETTYSSAPASFA
AOVMTNMAAALCLILGVGLPVIWIFQNMNLNGVGLMSAGRLDTPLGLVLP
HGLELTFVFAAGTGLRGLGLIDPGPRTRLAIEGPAALGMAVGLALVFGSA
IEGVTPLGTLRWAITTGVLALFLAVVYVYVCGAVRGETGVDEAARSAVTPVA
A"
complement(3489..3494)
3589..3595
3604..4560
/gene="SCE34.02"
3604..4560
/gene="SCE34.02"
/note="SCE34.02, possible integral membrane protein, len: 318 aa; similar to TR:069663 (EMBL:AL022121) Mycobacterium tuberculosis hypothetical 33.2 kD protein MTW025.043, 310 aa; fasta scores: opt: 734 z-score: 602.4 E(): 4.3e-26; 42.9% identity in 319 aa overlap. Contains possible hydrophobic membrane spanning regions"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein"
/protein_id="CAB88905.1"
/db_xref="GI:7649485"
/translation="MSEYTGENVALLERPARLPRLSRALAVLDLAVAVVAVTIAL
MYATSLDVAQNTASITATFVLYVGGTAVETLSRSGKMGACGRVYDDGGPIR
FRHSIVRGILGTYEITMTGLVACVSLVSARSRJSDVTRAGTLYVERPFSNAGE
PPEPPLAGRFSGLDLSAVDDLMILAVROVLARMGQIDPVGWMAARLADVAARTG
APVPEVPPAPVIAVLAIOEQRAERARAFEGASAEPTAAAVPAPAPAPAPAVPTAP
PVPAPAPAPAPPPSPGGLEVPDPVDRRRGTGFVPPA"
complement(4581..5195)
/gene="SCE34.03c"
complement(4581..5195)
/gene="SCE34.03c"
/note="SCE34.03c, unknown, len: 204 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SCE34.03c"
/protein_id="CAB88906.1"
/db_xref="GI:7649486"
/translation="MPGRGSLDHPDHPLEAEHFOCAPGSGMRYSLTTPGDA
OGSDVLAIDELGRPIRLLEHAAQVGAALIDGTWVTPPTGSLHAREGNARAFTG
TSPAFVAVARLRLTPSEATRVRLVALDPLVAPRTVQSMALLNREAHATNDPL
TYDEVQVTLDPGEOYAVHAGDVVLAAPGIELEDETPPSVFP"
complement(5306..6763)
/gene="sahH"
/note="SCE34.04c"
complement(5306..6763)
/gene="sahH"
/note="SCE34.04c, sahH, adenosylhomocysteinase, len: 485 aa; highly similar to SW:SAHH.TRIYA (EMBL:U40872) Trichomonas vaginalis adenosylhomocysteinase (EC 3.3.1.1) 0; 65.2% identity in 477 aa overlap. Contains pfam match to entry PF00670 AdoHcyase, S-adenosyl-L-homocysteine hydrolase and matches to Prosite entries PS00738 S-adenosyl-L-homocysteine hydrolase signature 1 and PS00739 S-adenosyl-L-homocysteine hydrolase signature 2"
/codon_start=1
/transl_table=11
/product="adenosylhomocysteinase"
/protein_id="CAB88907.1"
/db_xref="GI:7649487"
/translation="MTTVDRKDPKQVADLSAARGRKITLAHEHMPGIMAIRKEAYE
AOPLAARVTSGLHMYOTAVILLETVALGAQVWASCONISTODGAAATAVANGVT
PDPPQCPVFAFMGGETLEETWCKTEQALITPNPTPTGPNMLLDGDDATLTVHKEVY
EDKGVPSVTAASDEHRIVLELTTRVGEPPQWOTLASEIRGTELTGGVHLYTE
MHRDGTLLPFAIVNDAVTSKEDNKKYGRSLIDGINRAVDVILGKTAVVCGYGV

KGKAEIIRGCGARVITTEIDPICALQAMDFQVTTIDEVYKADIPVTTGNKDI
MAKDAKKHQAIVGNIGNHEDNEIDMAGIAPGPIYKDEKPOVHTTYPEDKVIYL
SEGRILNIGNATGHSFVMSNFPADDTLQIETFTPEDEPDTVYLPKHIDEVRL
HDSIGVKTITPEQADITGVKESGVPADHYR*
complement(5480..6733)
/gene="sahh"
/note="Pfam match to entry PF00670 AdoHcyase,
S-adenosyl-L-homocysteine hydrolase, score 1004.10,
E-value 3.3e-298"
complement(5921..5971)
/gene="sahh"
/note="PS00739 S-adenosyl-L-homocysteine hydrolase
signature 2"
complement(6479..6523)
/gene="sahh"
/note="PS00738 S-adenosyl-L-homocysteine hydrolase
signature 1"
complement(6773..6778)
/gene="SCE34.05c"
complement(7037..8017)
/gene="SCE34.05c"
complement(7037..8017)
/gene="SCE34.05c"
/note="SCE34.05c, possible transport protein, len: 326 aa;
similar to TR:AAF26917 (EMBL:AF210843) Sorangium
cellulosum putative transport protein, 305 aa; fasta
scores: opt: 629 z-score: 707.7 E(): 5.9e-32; 37.88
identity in 239 aa overlap. Contains Pfam match to entry
PF01545 Cation_efflux, Cation efflux family and possible
hydrophobic membrane spanning regions"
/codon_start=1

Query Match 44.8%; Score 655.2; DB 1; Length 37898;
Best Local Similarity 66.6%; Pred. No. 2.9e-140;
Matches 955; Conservative 0; Mismatches 473; Indels 6; Gaps 1;

QY 34 gagtacaagaattgcgacatacaactcattctctgcgcgttaagaacttaccctgct 93
Db 6739 GACTTCAAGTTCGCCGACCTCTCCCTGGCCGTTGGCCGCAAGATCACCTCTGCC 6680
QY 94 gagaagaagaatgcaggtctctatgcttctgtagagcgttatctcgtcttaagcattg 153
Db 6679 GAGCAGCAGATGCGCGGCTCATGCGGATCCGCAAGAGTACCGCCGACGCCCTC 6620
QY 154 aaggtgtcagaatcttggttccctccacatgacagtcagacagcgtctcatagag 213
Db 6619 GCCGGCGCCCGGTCACCGGCTCCCTGCATGACCGTGCAGACCGCGCTCATGAG 6560
QY 214 aaactcaagctctgtgtgtatgtacagatggtctcctcgaacatctctctacaaa 273
Db 6559 ACCCTGTCGCGCTGCGGCGCGGAGGTCGCTGCGCTCTGCAACATCTTCTCACCCAG 6500
QY 274 gatacagcgcgtgtgtctatgcttgcgccaacagcagacacagagaagccagct 333
Db 6499 GACCAACGCGGCGCGGCTATCCGCGCCGCAAGCGCCGCAACCGCCGCAAGGCG 6440
QY 334 atccagctcttcgccttggaagcgaaacatcccaagaaactactgggaacataaccgc 393
Db 6439 GTCCCGCTTCGCTGCTGAGAGGCGGACCTCTGAGAGTACTGAGTGTGTCAGAGACAG 6380
QY 394 gcttcacatgycagagatgltcaa-----ggcccaagcaggttctgtagtgcgt 447
Db 6379 GCGGTGACCTGCGCGAACCACCGCGCGCCGCAACATGATCTGCGAGCGAGGCT 6320
QY 448 gatgtacacctccatctccaaggtcttgatccgaacagcagcgtgtcttcagag 507
Db 6319 GAGCCCAACCTCTCTGTCACAAAGGCGCTGAGTACGAGAAAGCAGCAAGTCTCCCTCG 6260
QY 508 ccaacagaagctgaacactgaacacgcgtgtcttctgtaacatcaagcaggtcttc 567
Db 6259 GTGCAACCGCGGAGTGCAGACGACCGCGTCACTCTGCAACTCTCACCCGACGCTC 6200
QY 568 aaccaagaagaacactggtcacacagttgctgcgcgcatgaacggtgttcagagag 627

Db 6199 GCGGACAGCCCGCAGAAAGTGCACCCAGCTGGCGTCCAGATCCCGCGGTCTCACCGAGAG 6140
QY 628 acaacaacaggtgtgccaccgcctctaccagctcgaagaaggaaggaactctctccca 687
Db 6139 ACCACGACCGCGCTCCACCGCTCTAGAGATGCACCGGAGAGGACCCCTGCTGTCCGG 6080
QY 688 gccataacgttcaacagcgcgtgttacaagaattccaaagttcagatacaatctaggtgcgc 747
Db 6079 GCGATCAACGTCAACGACGCGGTCAACAGTCAAGTTTCGACCAAGATGACGCTGCCG 6020
QY 748 cactcccttactgattatcaaacgcgtctccagatctgcatatgattggcggaagaagct 807
Db 6019 CACTCCCTGATTCGACGCGCATACCGCGCACCGAGTCTGATCGCGGCAAGACCGCG 5960
QY 808 ctgcgtcgttgcggtacagcgcgtgtgcgaagggtcgcgtcctccctccgttgccaagc 867
Db 5959 GTCGTCTGCGGCTACGCGGAGTGCAGAGGCGTCCGCGGAGTCCCTGCGCGGCGAGGCG 5900
QY 868 gtcgcggtatctacatacaagaactcgaacccaatctgcgtctccagcgtgcgtggaagc 927
Db 5899 GCGCGGTGATCATCAACGAGATGACCCGATCTGCGCTGCAAGCGCGCGGCGATGAGCGG 5840
QY 928 taccagctgcgcgcgcacagaagaagtcgttaagaagatgctcatatcttgattatgacaga 987
Db 5839 TTCCAGGTTCACACGCTGAGAGAGTCTGCAACAGCGCGCATCTCTCCACACGAGCC 5780
QY 988 ggaacatgcatalcatctctgttgaacatgtagcccgatgaagaagatagatctgttc 1047
Db 5779 GGCACACAGACATCATCATGAGGCCAAGGACATGCGCAAGATGACACACGCGCATGCTC 5720
QY 1048 ggttaacatgcgcacactcgaatacgaatgtatataagaatggtcctatgaataccagcgc 1107
Db 5719 GGCACATCTGCGCCACTTCGACACAGATGACATGCGCGCGCTGCGCACACCGCGGCG 5660
QY 1108 atcaagacatctcccaatcgaagaagatgacatgttggaattccagatggtccagct 1167
Db 5659 ATGCTCAAGACGACGATGCAAGCCGAGGCTCCACACCTTGAGCTTCCGACGCGAGGTG 5600
QY 1168 atccctctctgtcgtgagcgctcttcaacttgcgtcgtcgtacagatccatct 1227
Db 5599 CTCATGCTCTCTCCGAGGCGCGCTGCTGTAACCTGGGCAACGCCACCGCGCACCGCTG 5540
QY 1228 ttggtatgtcaatgtatctcaacaacagaacactcgtcgtcgtacgttgcgttgcgaaga 1287
Db 5539 TTCTGATATGTCACACTCTTCGCGGACGACACCTGGCGGCAAGTCAAGCTTTCACCAAG 5480
QY 1288 agaggaatctcgaagaagaaggtttacacacttccgaagcattcgtatgaagaatgct 1347
Db 5479 CCGAGAGAGTACCCGACCGACGCTTACGTGCTGCCCAACGACCTTGAGAGAGGTGCGC 5420
QY 1348 cgcctccactcgcgtatctcgtatgtccaccttcaaaagcttacaagaagaagcgtgac 1407
Db 5419 CGGCTCACCTCGACTCCCTTCGCGGTCAAGTGCACAGTCCGCGCGGAGAGCGGAGC 5360
QY 1408 tacatcaagttcagttgaggttccttacaagtcgtatgcttaccgttataa 1461
Db 5359 TACATCGGCTCAAGGTGAGGCGCCCTTACAAAGCGGACCACTACCGCTACTGA 5306

RESULT 4
SAR416377 3800 bp DNA linear BCT 15-OCT-2001
LOCUS SAR416377
DEFINITION Streptomyces argillaceus mtnz gene, mtna gene and mtnh gene.
ACCESSION AJ416377
VERSION AJ416377.1 GI:16197721
KEYWORDS adohcyase; AdoMet synthetase; MTHF reductase; mtna gene; mtnh gene;
mtnz gene; thioesterase.
SOURCE Streptomyces argillaceus.
ORGANISM Streptomyces argillaceus
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 3800)
Fernandez Lozano, M., Lombardo, F., Brana, A.F., Salas, J.A. and Mendez, C.

TITLE Generation of a Streptomyces argillaceus mutant by deleting two methyl cycle-related genes located within the mithramycin gene cluster leads to higher production of mithramycin

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3800)

AUTHORS Mendez, C.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-2001) Mendez C., Biología Funcional e IUOPA, Universidad de Oviedo, C/ Julian Claveria s/n, 33006 Oviedo, SPAIN

FEATURES

source Location/Qualifiers

1..3800

gene /organism="Streptomyces argillaceus"

CDS /db_xref="taxon:41951"

7..789

/gene="mtm2"

/function="putative role in polyketide chain biosynthesis and release"

/codon_start=1

/transl_table=11

/product="putative thioesterase"

/protein_id="CAC94888.1"

/db_xref="GI:16197722"

/translation="MAAOPRLCYCFANAGACVSAPFGMAARVPDVPVLLPGND RRREPRVTTRAAALDDVRLFGAPTAPEVLVIGALVATVGRALHEAGRGPA LLAAGCPPEPLPEPPAYALGPDALERTLDLMAVPSGTGTEAGIGIDRAVPLV RDOLTLARALACTGSSPADLITPLILVADDDPMASPAVAGSMSEGTPLHIT VSGGHFVRGPELRLGACALAPARTTPAAGVSTGN"

918..2300

/gene="mtm1"

CDS /db_xref="taxon:41951"

918..2300

/gene="mtm1"

/function="putative role in an activated methyl cycle"

/codon_start=1

/transl_table=11

/product="Adomet synthetase and MTHF reductase fusion"

/protein_id="CAC94889.1"

/db_xref="GI:16197723"

/translation="MSGHFTSESVEGHPDKIADISGTPIDALLREDPRSHAVET LITGOVAVADKVTGAYVPIAOLVRDAVLTIGDSAGPDGACGVSIGPSOD IAQGVDRAYESLVEBAAGADIDELDROAGCGIMECATDESPLMPLPITLAVL AKRLIDVRNGDLPPTFARTARPGSPEDRAGRARVIMDTIRRGALRPDESVTEAG SPDRIVVALTRIAETRLRPAHLVAGHSAVELKQIIGRADVGIIDVYALRGDPD GDPRGPMSPHEGFEYVAHELVEPVSIGFTIRMAFPRHPSPMESDIIRHVAC RAGADVATTOGFEFGDGLRLDRPAAAGCHPIPIEIPAMOVROIIRPAELSAVF TEDLAHRELOAGNPNVGHRIQVDHATAEERLALAGVPGIHCITLNRSPALDIHRL LTRSRSLSVR"

2290..3738

/gene="mtm1"

CDS /db_xref="taxon:41951"

2290..3738

/gene="mtm1"

/function="putative role in an activated methyl cycle"

/codon_start=1

/transl_table=11

/product="AdoHcyase"

/protein_id="CAC94890.1"

/db_xref="GI:16197724"

/translation="MSADPTDFKVDLSAERGRKEITLAEHMPGLMSIRREYAEAO PLACRVATGSLHMTYOTAVLITETVALAGOVASINISSTOGHAAATVAGPGTDP NPAVPYPAKGEITLIEYWCITEGALTWDSPISGGMILIDSDGATLLVHKVEYRK DGKVPVDAIVNDENRVLDELNRTITDSQKWTQSLRGVTEETTVHRIENKO RDGVLPFAINVDENRVLDELNRTITDSQKWTQSLRGVTEETTVHRIENKO AARSJGOGARIVTEIDPICALQAMDPQYVTTDEVDKADIFITTKNKDIIMAA DMARKHQAIVGNIGHDEIDMAGLAKVGYKDEIKPOVHTWTFPDGKKIIVLSGS RLINLGNATGHSPFVMSNFADOTLAQITLPIKPGDYPIGVYTLPHKLDKVARLHLD ALGVKTLTLRPEQASLYGVNDGPKYKSHYR"

BASE COUNT 590 a 1485 c 1197 g 528 t

ORIGIN

Query Match 42.1%; Score 615.2; DB 1; Length 3800;

Best Local Similarity 65.3%; Pred. No. 4.2e-131;

Matches 937; Conservative 0; Mismatches 488; Indels 9; Gaps 2;

QY 34 gagtacagattgacgacatcaaccatccatgcttcctggcgtaaggaacttaccctgct 93

Db 2308 GACTTCAAGGTGCGCATCTCTCTCTGCGGAGTGGCGGCAAGGATCAACCTCGCC 2367

QY 94 gagaaggaatgacaggtcttactatgctctcgttagcgatttaccgcttcaagcaatg 153

Db 2368 GAGCAGAGATGCGCCGCTGATGTCGATCCGCGAGGAGTACGCGGAGCCGACCCCTC 2427

QY 154 aagagtgtcaagaatctctgcttccctcccaatgacagatccagaagccgcttcatgag 213

Db 2428 GCCGCGCGCGGTCACCGGCTCTCCGACATGACCGTGGACACCGCGCTCTCATCGAG 2487

QY 214 acattcaacgctctgtgtctgtagtgcagatggcttcgcaacatctctctacacaa 273

Db 2488 ACCCTGCTGCGCTGCGCGCCGACAGTCCGCTGGGCTCTCTGCAACATCTTCTCCACG 2547

QY 274 gataagcgcgtgtgtctatgctgtgtgcccacaagacacacagaagaagccagccgt 333

Db 2548 GGCACGCGCGCGCGCCGACATCGCGTGGCGCCGACGCGGACGCGCGCGCGCGC 2607

QY 334 atcccaagcttcgcttggaagcgcaaacatcccaaatatcttggaagaacacatacgc 393

Db 2608 GTCCGCGTCTGCGCTGGAAGGCGGAGACGCTGGAGAGTACGTGGTGCACGAGCAG 2667

QY 394 gctctacatgagcagatgctc-----aaggcccaagcaggtgtgctgtagtgtgt 447

Db 2668 GCGCTACCTGCGCGGAGACGCCACCGCGGCGCGGACGCGGCGCGGCGCGGCGGCT 2727

QY 448 gatgtacacatccatctcccaagggctcgaatctgaatgaacaagccggtgtgtccagag 507

Db 2728 GAGCGCACCTTGCTGTGCTCCCAAGGGGCTGTGAGTACGAGAGGAGGAGGAGTCCCTCG 2787

QY 508 ccaacagaagctgacacacccgaataccgctcgtgtctgtgtctacacgaagctgtc 567

Db 2788 GTTCGACACCGCGGAGAGACGACGACCGGCTCTCTGCAATCTTCAACCGCACATC 2847

QY 568 aaccaagaagaacacacatgacacacagctgtgtcggcgatgaacggtgtgtccagaag 627

Db 2848 ACCGAGCGGCTCCCAAGATGAGACGACGACGCGCTCTCTGCAATCTTCAACCGCACATC 2907

QY 628 acaagaacaggtgttccacgctctacacgctcgtgagaagggcgaaactcctctccca 687

Db 2908 ACCACACCGCGGCTCCACCGCTGTACGAGATGAGGCGGCGGCTCTGTGTCCCG 2967

QY 688 gccatcaacgtaacagacgctgttacaaggtccaaatcgaatacatcgaatcgcgcgc 747

Db 2968 GCGATCAAGCTCAACGACGCGGCTACCAAGTCCGAAGTTCGACACMAATACGCGTCCCG 3027

QY 748 cactccctatcgaatgatacaacgctgtcgaatgataatcgaatcgcgcgaagaagct 807

Db 3028 CACTCGCTGATCGACGCGCATCAACCGCGGACGCGGCTCTGATGCGGCGGCGGCGGCG 3087

QY 808 ctgcgtatggtgttaagcgatgttcgcaagggcgctgaatccctcgtggtgcaagcg 867

Db 3088 GGTGTTTGGCTACGCGGCGGCTGCGGCAAGG---CTGGCGGAGTCTCGGCGGCGGCG 3144

QY 868 gctcggttatacatcaagaagactgaccacaaatcgtcgtctcccaagctgcaatggaagcg 927

Db 3145 GCCCGGCTATCGTCAACCGAGATGACCGCATCTGCGGACGCGGCGGCGGCGGCGGCGGCG 3204

QY 928 taacaggtccgacgcatcgagaaggtcgtcaaggaatgtaatactctgttaacatgaca 987

Db 3205 TACCAAGTCAACGACCTCGAGAGTGTGTCGACAAAGCGCGGACATCTTATCATCACGAGCC 3264

QY 988 ggaacatgagatacatctctgttgacatgataggccgaagatgaagagataagctatgtc 1047

Db 3265 GGCACAAAGGACATCATCATGCGCGGAGACATGGCGCAGAGTACACCAAGCCATCTGTC 3324

QY 1048 gtaacatcgacacatcgaataagaaatgatacagatgagctcgaataataacacagcg 1107

Db 3325 GGAACATCGCGCACCTTCAACGAGATGACATGCGCGGCGCTGCGCAAGGTCGCCCGGCG 3384

QY 1172 tccctctgtgagggcgccctcttctaactgtgctgcgtacaggtcaacccatcttgc 1231
 Db 1248 TCGTCTCCGAAGGTGGCTGTGGAACCTTGGCAACGCCGACCCATCATTTG 1307
 QY 1232 ttatgtcaatgtatctcaaacacagacacgtcagctcagctcagccctcaagaagaag 1291
 Db 1308 TCAATGTCAACTCTTTGCCCATCAGACCATTTGCCGAGATCGAAGCTTCCAAACGAC 1367
 QY 1292 gaaatctcgagaagaagtttaacacatctcgaaagcatctcgatgaagaagtcgtcc 1351
 Db 1368 GACAGTACGAAGACGAGGTCTACCTTCCCTTAAGGTTCTCGACGAAGAGTGGACGCA 1427
 QY 1352 tcaacctcgatctcgtatgctgcagcagctcaagaagcttcaagaagaagcagctac 1411
 Db 1428 TCCAGTTGAGAGCTCTCGGGGTCACTACCGAATGACCAAGAGAGAGGCTGAGTACA 1487
 QY 1412 tcaacgtctcaggtgaggtccttaagaatctgattgcttacttaaa 1461
 Db 1488 TCGGCGTTGAGAGTTGACAGGCCCATTCACCGGACGACACTACCGCTACTAA 1537

RESULT 6
 AX244105 1557 bp DNA linear PAT 28-SEP-2001
 LOCUS AX244105
 DEFINITION Sequence 97 from Patent WO0166573.
 ACCESSION AX244105
 VERSION AX244105.1 GI:15859169
 KEYWORDS
 SOURCE Corynebacterium glutamicum.
 ORGANISM Corynebacterium glutamicum
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacteriales; Corynebacteriaceae;
 Corynebacterium
 1 (bases 1 to 1557)
 Pompejus, M., Kroege, B., Schroeder, H., Zelder, O., Habermeyer, G.,
 Kim, J. W., Lee, H. S. and Hwang, B. J.
 Corynebacterium glutamicum genes encoding metabolic pathway

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 BASF AKTINSELSCHAFT (DE)
 Location/Qualifiers

FEATURES
 source
 CDS
 1..1557
 /organism="Corynebacterium glutamicum"
 /db_xref="taxon:1718"
 101..1537
 /note="RXN00132"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC88531.1"
 /db_xref="GI:15859170"
 /translation="MAOYDFKVAADLSLAEGRHQIRLAIEYEMPGILMLKREFADEOP
 LKARIASIHMTVOTAVLIEITLALGAVERMASCIFSTODEAAALVIGSGVEEP
 AGVYFPAKGBESLEBYWNCINIFSGWDELPMIILDDGDAITMAVIRGREOAGIVP
 PAEANDSDETIAFLMLREVLAIEPGRKRIAAVAVYETTTGVRILHFAEGYL
 PEPANNDVAVTKSFNDKRYGRHSLLDGINRAITDMGKNVLCVGDGKCAEA
 FDDGARVAVTEADINALQALMDGYSVTVDAIDADIVITATGKDIISFGOMK
 MKDHALLGNIGHDEIDIMHSLHSDVYRTITKPYDETFSTGRSIIYLSERLLN
 LGNATGHSFVSNSEFADOTIAOIELFONEGYENEYRILPKYLDKVARLHVALGS
 QLTETLEOAEIIGYDVAGPFRKPEHYR"
 BASE COUNT 348 a 440 c 431 g 338 t
 ORIGIN

Query Match 40.8%; Score 596.4; DB 6; Length 1557;
 Best Local Similarity 64.5%; Pred. No. 8.5e-127;
 Matches 923; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 32 tcaagatcagaatctcgacatcaacctcatgttctcgscgtaagaactacccttg 91
 Db 114 TGGACTTCAAGGTGGCCATTCTTCACTAGCAGGACGACGCTACACAGATTGCTTGG 173
 QY 92 ctgagaagaagaatgcaggtctctatggttctctgtgagcgcttattcgcctcctaagcat 151

Db 174 CAGAGTATGATGATGCCAGGCTCATGACAGTTGGCAAGAAATTCGACAGACGACCTT 233
 QY 152 tgaaggtgtcagaatctctgttccctccacatgaagttcagaagccgctccatcgc 211
 Db 234 TGAAGGGCCCCGGAATTTGGTGGTTATCCACATGACAGGTTCACACCGCGCTTATG 293
 QY 212 agacacacagctctgtgtgtcgtatgtcagatgagcttcctcgaacatctctacac 271
 Db 294 AGACCTTCACTGCTTTGGGGCGCTGAGTTGGTTGGCTTCTCTGACATTTTTCACACC 353
 QY 272 aagatacagccgtgtgtctatcgttctgtcggccaaagagcaacacagaagccagccg 331
 Db 354 AGGATGAGCGTCAAGCGGCTATGTTGTCGGC---TCCGGACCCGTGMAAGCAGCTG 410
 QY 332 gttatccagatcttgcgtcgtgaaggggaaacacccacagatcttggaagacatcac 391
 Db 411 GTGTTCCAGTATTTGGGTGGAAGGGTGAAGTCACTGGAGAGTACTGCTGGTGCATCAACC 470
 QY 392 ggcgtcacaatggtccagatgtgtcaagggccacagcaggtgtcgtatgtgtgtga 451
 Db 471 AGATCTTCAAGTGTGGGGCGATGAGC---TGCMAAATGATCTCTGACGACGGGGTGA 527
 QY 452 ctacactcctcatctccaagggtcctcgaatctgaagaacagccggtgtgtccagaagca 511
 Db 528 CCACCATGGCTGTATTTCCGGGTGCGGAATACGAGAGCGGTGTGTTGCCACACGAG 587
 QY 512 cagaagctgacaacccctcgaatcagctgcgttctgtctacactcaagaaggtcttcaac 571
 Db 588 AGGCCACAGATTTCCGATGATGATACATCCATTTCTTGGGCACTGCGTGAAGTTCTTGG 647
 QY 572 aagaacaagaacacatgtgcaacaggtgtcgtcggcatgaacggtgttccgaagaaca 631
 Db 648 CAGAGCTTGCAAGTGGGGGCAAGATCGCTGAGCGCTTAAAGGTGTCACCGAGAAACA 707
 QY 632 caacaggtgtcaccgctcttaccagctcgaagggaggggaaactcctctccagacca 691
 Db 708 CCACCGGTGACACCGCGCTTACCATCTTCCGTGAAGAGGCGTGTGCTTCCAGGCA 767
 QY 692 tcaactcaacagcgtcgttcaacaaagttccaaatctcgaatacatcagctgtccgacat 751
 Db 768 TGAAGCTCAACGACGCTGTACCAACAGTCAAGTTGATTAACAGTACGGACCGCACT 827
 QY 752 cccatcagatgtatcaacacggtgtcctcgtatgtcatgtatcgtgcggcaagacgtctgc 811
 Db 828 CCTGTATCGACGGCATCAACCGCGCACTGCAATGCTCATGCGCGCAAGAACGTCCTG 887
 QY 812 tcaaggttcaacggtcgtatgtcgaagggcgtcgtcgaatccctcgtgtgcaagcgctc 871
 Db 888 TCTGCGGTTACGGCGATGTGGCAAGGGCTGCTGAGGCTTTCACGCGCGGCGCTC 947
 QY 872 ggcgtatcaacagaactcgaacccaatctgtgcgtctcagagctgtcgtatgaaggtctac 931
 Db 948 GCGTCAAGGTCAACGAGGTGACCAACCAATCAACGCTTCAAGGCTGTGATGATGCTACT 1007
 QY 932 aggttcgcgcgcatcgaaggaagtcgtcaaggtgtcgtatctcgttctcgtatcgaagga 991
 Db 1008 CTGTGTGTCACCGTTGATGAGCGCATTCAGACGGCGGAATGCTGATCAACCGGACGGGA 1067
 QY 992 actgcatacatctcgttgcatagtatggtccagatgaaggaataaggtatgttcgtga 1051
 Db 1068 ACAAGGACATCATTTCTTCTTGAGAGCATGCTCAAGATGAAGATGACGCTGTGCGGCA 1127
 QY 1052 acatcggcacttcogatacagaatgtatcagatgagctcatgaataaccagacatca 1111
 Db 1128 ACATCGGTCACTTTGATTAAGAGATGATGATGATTCCTCTTGGACCGCGACGAGCTGA 1187
 QY 1112 agacatcccaatcaacagcgaatgaagcagatgtggaatctccagatgcaagcgtcatc 1171
 Db 1188 CCGGACACAGATTAACCAAGTGTGACGAGTTACCTTCCACCGGTGCTCATATA 1247
 QY 1172 tccctctgtgagggcgccctcttctaactgtgctgcgtcgtacaggtcaacccatcttgc 1231

RESULT 8
 AX120920 1422 bp DNA linear PAT 11-MAY-2001
 LOCUS Sequence 836 from Patent EP108790.
 DEFINITION AX120920
 ACCESSION AX120920
 VERSION AX120920.1 GI:14037635
 KEYWORDS
 ORGANISM
 SOURCE
 Corynebacterium glutamicum.
 Corynebacterium glutamicum
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 Corynebacterium.
 REFERENCE
 1 (bases 1 to 1422)
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 836 20-JUN-2001;
 KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source
 location/Qualifiers
 1..1422
 /organism="Corynebacterium glutamicum"
 /db_xref="taxon:1718"
 BASE COUNT 309 a 414 c 401 g 298 t
 ORIGIN

Query Match 40.7%; Score 594; DB 6; Length 1422;
 Best Local Similarity 64.5%; Pred. No. 3e-126;
 Matches 920; Conservative 0; Mismatches 500; Indels 6; Gaps 2;

32 tcgaagacagaatttcgcagacatcaacatcctcagctcgcgtaagaaacttaaccttg 91
 11
 Db 2 TGGACTTCAAGGTTGCGCATCTTTACTAGACAGAGGACGATCAGATTGCTGTG 61
 11
 Qy 92 ctgagaagaaatgcgaagctcttaagttcctcgtgtagcgcttaccgctcttaagcaat 151
 11
 Db 62 CAGAGTATGAGATGCGAGGCTCATGCTGCGCAAGGATTCGCGAGAGAGCAGCCTT 121
 11
 Qy 152 tgaagagtgtagaattcgttccctccacatgacatgcagaaagcgcgtctcactcg 211
 11
 Db 122 TGAAGGGGCGCCGAATGCTGTTATTCACATGACGATGCTCCAGACCGCGCTTATTC 181
 11
 Qy 212 agaacatcagactcttgctgctgaagtcagatgagtggtctctcgaacatctcttaac 271
 11
 Db 182 AGACCTTACTGCTTTGGCGGCTGAGGTTGCTTGGGCTTCTCGACATTTTCTCACCC 241
 11
 Qy 272 aagataagcgcgtcgtctatcgttctggtgcccacagacacccagagaagccagccg 331
 11
 Db 242 AGGATGAGGCTGACGCGCTATCTGTTGTGGC--TCCGGCACCGCTGGAAGAGCCAGCTG 298
 11
 Qy 332 gttaccagctcttcgctggaagagcgaaacactcccaataacttggaagaaacatacc 391
 11
 Db 299 GTGTTCCAGTATTCGGGTGGAAGGTGAGTACATGAGAGGTACTGGTGGTGCATTCACAC 358
 11
 Qy 392 ggcgtctcaacatgacagatgltcaagccacagcagatgttcgatgatggtgtagatg 451
 11
 Db 359 AGATCTTCAAGTGGGGCGATGAGCT--GCCAAACATGATCTCCGACGACGCGGTGACG 415
 11
 Qy 452 ctacactcctcaatctcccaaggggttcgaatgaaacagcgcgtgctgctcagagcaaa 511
 11
 Db 416 CCACCATGGCTGTATTCGGCGGTCGGAAATACAGACAGGCTGTGCTGTGTTCCACGACAG 475
 11
 Qy 512 cagaagcttgacaacactcgaaataaccgtgcttctgtaactcaacgaagtlctcaacc 571
 11
 Db 476 AGGCCAAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 535
 11
 Qy 572 aagacaagaacacttgacagatgctgctgctgctgctgctgctgctgctgctgctgctg 631
 11
 Db 536 CAGACCTTGGCAAGTGGGGCGAAGATGCTGAGGCGCGTTAAGGGGTGTACACCAAGGAACA 595
 11
 Qy 632 caacagtgctcaccgcctctacagctctgagagaggggcaaacctctctccagcga 691

Db 596 CCACCGGTGTGACCGCCTGTACCTTCCCTGGAAGAAGCGTTCCTTTCCACGAGA 655
 11
 Qy 692 tcaagctcaacgaagctgttacaagttccaaattcgaatacactcaagctgcgcgaact 751
 11
 Db 656 TGAAGGTCAACGACGCTGTACACCAAGTTTGAACAACTAGCGGCCGCCGCTCT 715
 11
 Qy 752 ccctacatgctgatacaacacgctgctcgtatgcatatgactcgcgcgaagaacgctcgc 811
 11
 Db 716 CCCTGATGACGCGATCAACCGCGGCACTGATGCTCATGAGGGCGGCAAGAACGTGTG 775
 11
 Qy 812 tcaatggttaagggatgctgcgaagggctgctgctcaatccctcgtggtgccaagcgctc 871
 11
 Db 776 TCTGGCGTTACGCGCATGTCGCGAAGGCGCGCTGAGGCTTTGACGCGCGCGCTC 835
 11
 Qy 872 ggtatatacaccagaacatccgacccaatctgcgctcccaagctgcgcgtcgaaggaactacc 931
 11
 Db 836 GCGTCAAGTCAACCGAAGCTGACCCCAATCAACGCTTTCAAGCTCTGATGATGATGATG 895
 11
 Qy 932 aggtccgcgcacacgaagagtcgtcgaagatgctgatactcgttacaatgacagaaga 991
 11
 Db 896 CTTGGGTCAACGTTGATGAGGCGCATGAGGACGCGCATGCTGATCACCGGACGCGCA 955
 11
 Qy 992 actgcatatcactctgttgcacatgtagtgcacagatggaagataagctatgtcgta 1051
 11
 Db 956 ACAAGGACATCATTTCTTCCGACGACATGCTCAAGATGAAAGATCACGCTGCTGGGCA 1015
 11
 Qy 1052 acatcgccactcgtatgaacgaattgatacagaatgctcgtcctcaatgaataacagagcatca 1111
 11
 Db 1016 ACATCGGTCACTTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1075
 11
 Qy 1112 agcaatcccaatcaagccagaataagacatgctggaattccacagatgccaagctatcc 1171
 11
 Db 1076 CCCGACCAAGATCAAGCCACAGGTGACGAGTTCACCTTCCACCGGTGCTCATCA 1135
 11
 Qy 1172 tccctcttgtagagggccgctcttcaacttgctgctgctgctgctgctgctgctgctgctg 1231
 11
 Db 1136 TCGTCTGTCGAAGTTCGCTGTTGAACCTTGGCAACGCCACCGACGACCATATTTG 1195
 11
 Qy 1232 ttatgcaatgltatcaacaacacgaactcgtcagctcagctcagctcagctcagctcag 1291
 11
 Db 1196 TCATGTCGAACCTTTTCGCCGATGACATGTCGCGAGATGATGATGATGATGATGATG 1255
 11
 Qy 1292 gaaatctcgaagaagaggtttacaacactccgaagatctcgtatgagaagaatgctgcgc 1351
 11
 Db 1256 GACAGTACGGAACGAGAGTGTACGCTGTGCTTAAGTTCTTGACCAAAAAGGTGACACGCA 1315
 11
 Qy 1352 tccactcgcgactctcgtatgctcacttacaagacttacaacagaagcagctgactaca 1411
 11
 Db 1316 TCCACGTTGAGGCTCTTGCGGGGTCAAGCTCAACGAACTGACCAAGAGCAGGCTGATGACA 1375
 11
 Qy 1412 tcaagctcgaattgaggttcccttaagaagctgactgactgactgactgactgactg 1457
 11
 Db 1376 TCGGCGTTGACGTTGACGCGCCATTCAAAGCGGAGCACTACCGCTCA 1421
 11

RESULT 9
 AF262755 1488 bp DNA linear BCT 30-MAY-2000
 LOCUS Mycobacterium bovis S-adenosyl-L-homocysteine hydrolase gene,
 DEFINITION complete cds.
 ACCESSION AF262755
 VERSION AF262755.1 GI:8101959
 KEYWORDS
 SOURCE
 ORGANISM
 Mycobacterium bovis.
 Mycobacterium bovis
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;
 Mycobacterium; Mycobacterium tuberculosis complex.
 REFERENCE
 1 (bases 1 to 1488)
 AUTHORS Pawar, S.N. and Nayak, R.
 TITLE Complete genomic DNA sequence of the S-adenosyl-L-homocysteine
 hydrolase gene from Mycobacterium bovis BCG

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1488)
AUTHORS Pawar, S.N. and Nayak, R.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2000) Microbiology and Cell Biology, Indian
Institute of Science, Malleshwaram, Bangalore, Karnataka 560012,
India

FEATURES

source

CDS

1.1488
/organism="Mycobacterium bovis"
/strain="BCG-Pasteur"
/db_xref="taxon:1765"
1.1488
/codon_start=1
/transl_table=11
/product="S-adenosyl-L-homocysteine hydrolase"
/protein_id="AAF72670.1"
/db_xref="GI:8101960"
/translation="MTGNLVTNRNSLTDPVRNGIDFKIADLSTADFGKREIRIAEHNP
GIMSLRREAYOPLKGRIRISSLSIMTVOTAVLITLALGAEVNMASTIFSTODHA
AAAVVGGHGPDEPKVPEVFAWKGETLEIYVAAEQMLTPDPDPANMILDDGDA
TMLVLRGMYEKAGVVPABEDDPAEWKEVFLMLRTREFDKKWKRIASVGTVEE
KTLIVLRLOFPAAGDLAPALINVDSTKSKFDNKGTRHSLIDGINRTDLIGK
KVLICGVDVGKCEAKKGAGGARSVTEIDPILNLAQMEHSDVYVEAIGDADIV
VTATGNKDIIMLEHKKAKMDHAILGNIGHFQNEIDMAGIERSATVNVNRPQDUMTF
GTPRSSTIVSEGRILNIGMATGHPSEFVANSFPANOTIAOIELMTINDEYDNEYTRLR
KHIDEKARIVEALGHLTKTKSOAEYLGVDEGPKPDHYR"

BASE COUNT 313 a 475 c 474 g 226 t

Query Match 36.6%; Score 535.4; DB 1; Length 1488;
Best Local Similarity 61.3%; Pred. No. 9,2e-113;
Matches 879; Conservative 0; Mismatches 551; Indels 3; Gaps 1;

QY 32 tcgagtaacagattgcgcacatcaactcattctcgtcgcgtaaggaacttaaccttg 91
DB 56 TCGACTTAAATGCGCGACACTGTGACTAGCGGATTTCGGCCGAAGAAGACTCGGATCG 115
QY 92 ctgagaagaatgcgaagttctatgtctctcgtcgtcgaagcgttattccgcttcaagccat 151
DB 116 CCGAGCAAGAGATGCGCGCTGATGTCGTCGCGCGAGTATGCCAGAGTGCAACCCC 175
QY 152 tgaaggtgttagaatctcgttccctccatcatgacatccgaacagccgtctcctatcg 211
DB 176 TGAAGGAGGCCCGGATGCTGGTGGTGCACATGACGGTGCACCGCGCTGTGTTCG 235
QY 212 agaacactcaagctctgtgtgtgtatgtatgagtgagcttcctccatcatcttctaacac 271
DB 236 AAGCCCTACCGCGCTGGGCGCGCGAGTCCGTGGGCTTCGTGACACATTTCTCCACC 295
QY 272 aagataagccgctgtgtctatgttctgtcgcgaacagcgaacaggaaggaagccg 331
DB 296 AGGATCACGCGCGCGCGCGCTGTGTCGCGCCGACGCGACCCCGAGAGAGCCCAAG 355
QY 332 gttaccagctctcgtcgtggaagcggaacacatcccaataacttgggaacacatacc 391
DB 356 GTGTCCGGTGTGGCTGGAAGGCGAGACCTCGAAGAGATCTGTGGCGCGCCAGC 415
QY 392 ggcgtctcaatgagcaaga-----tggtaagcgcccaagcaggtgttcgatgatgtgtg 448
DB 416 AGATGCTCACTGGCGGAGACCCGACAGCCGCGCAACATGATCTCGATGAGCGCGGTG 475
QY 449 atgtactactctcatctccaaggcttcgaatctgnaaacagccggtgtgttctcagaagc 508
DB 476 ACGGCACCATGTTGTGTGCGCGCATGACATGAGAAAGCGCGCGGTGTCGCCCGC 535
QY 509 caacagaagctgacaactgcaactgcaactcgctgtctctgtctacacactcaagcaggtctca 568
DB 536 CCGAGGAGAGACACCCCGCGAGTGAAGGTTTCTCTGAACCTGTAACGAGACCCCGTTCG 595
QY 569 accaagaacaagaacactgtgacacagttgtcgcgcatgaacggtgtttccgaagaga 628

DB 596 AGACCGAAGAGACMACTGAGACCAAGATAGCCGAGTCCGTCAGAGGCGCTCACCGAGAGA 655
QY 629 caacacaggtgtccacgcctctaccagcctcgaaggaaggaaggaactcctctccag 688
DB 656 CCAACACCGCGCGTCTCGGCTCTACCAATTCGCCGCGCGGAGATCGGCTTCCCG 715
QY 689 ccatcaagctcaacagcagctgttacaagagtcgaagtcgaatctcagctgtccg 748
DB 716 CGATCAACGTCACAGCATCGGTGACCAAGTCAATTCAGACAAAGTACGCGACCTCG 775
QY 749 accctctatcgtatgtatcaacagctgtctccgattgtatgatacggtggaagacgtc 808
DB 776 ACTCCCTGATCGAGGCGATACCCGCGCACCGCGCTGATTCGGCTAAGAGTCC 835
QY 809 tgcataaggttaccgagatgtcgcgaagcgtcgctcccaatccctccgttgcgaagcg 868
DB 836 TCATTCGCGCTACCGGAGAGTCCGTAAGGCTGTGCGGAGGCGATGAAGGCGAGG 895
QY 869 ctgcggttatacatcaagaactcgaccatctcgcgtctccagctgtccatgaagcgt 928
DB 896 CCGCGGCTTCCTCCACCGAGATCGACCGGATCAACGCGCTGACGATGATGAGGCT 955
QY 929 accaggtccgcatcgaaggaagtcgtcaagatgtcgtatcttcttatacgaagc 988
DB 956 TCACCTGCTACCGTCAAGAGGCGCATCGGGACCGCGACATCGTTAACCGCGACG 1015
QY 989 gaactgcgatacatctcgttgcataatgagccagatgaagataagcattgtcgt 1048
DB 1016 GCAACAAAGACATCATCTCTCGACGACANTTAAGCGATGAGGACACCGATCTGTG 1075
QY 1049 gtaacatcggcgaactcgaataagaaatgtatagatgagcctcgaataccagga 1108
DB 1076 GAATATCGGCGCTTCGACAGAGATGACATGAGCGCGGCTGAGCGCTCGGAGCA 1135
QY 1109 tcaagcattcccaatcaagcgaatgaacatggtggaattcccaatgcccagccta 1168
DB 1136 CAGGGTCAACGTCACAGCTTCAGTGCACCTGTGACCTTTGGGACACGCGCTCGA 1195
QY 1169 tctcctcttctgtcgaagcgccctcttaacctgtcgtcgaagcagtcaccatct 1228
DB 1196 TCATCTGCTGTCGAGGCGGCGGTGCTGAACCTGGGGAATGCCACCGGACCCCTCGT 1255
QY 1229 tcttatgtcaatgtatctcaacaacagacactcgtcagctcgaactctcagaaga 1288
DB 1256 TCGTATGAGCAACAGCTTCGTAACGACAGATCGCCCAATGACGCTGTGACCAAGA 1315
QY 1289 gaggaaatcgggaaggaaggtttacacacttccgaagcactcgtatgaagaagtcgtc 1348
DB 1316 ACAGCAGTACGACAAAGAGGTGTACCGGCTGCCCAAGCACCTCGACGAAGAAGTGGCTC 1375
QY 1349 gctccacactcgaatctcgtatgtccaccttaacaaacttacaagaagcaggtgtact 1408
DB 1376 GAATCATGTGAGGCGCTTGGGCGGTACCTGACCAACGCTCAACGAAGAAGCGCGAAT 1435
QY 1409 acatcaaggttccagttgaaggttccctacaagtcgtatgtaccgttataa 1461
DB 1436 ACCTCGCGTCAAGCGTGAAGGTCCTCAACAAAGCGGACCACTACCGCTACTGA 1488

RESULT 10
LOCUS AX023852 1488 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 23 from Patent WO0021983.
ACCESSION AX023852
VERSION AX023852.1 GI:10184196
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriales; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE
AUTHORS Skjot, R.L., Andersen, P., Rosenkrands, I., Weidling, K., Okkels, L.M.,

CDS complement(7620..9107)
 /gene="MT3346"
 /note="similar to SP:P51540 PID:1118009; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="adenosylhomocysteinase"
 /protein_id="AAK47688.1"
 /db_xref="GI:13883166"
 /translation="MTGNLVTKNSLTTPDYRNGIDFKIADISLADFGKRLRIAREHEMP
 GIMSLRREYAEVQPLKGLRISGLIHTVOTAVLIEITLALGAELRMAASCNIFSDHA
 AAAYVGHPTDEPKGVVPAFMKGETLEBYWMAADMLTWPDPKRNMLIDGGGA
 TMLVLMQYKAYGVVPAEEDDPAEMKVELNLRLRFETDKWTKRIASGVTEB
 TTGVLRLYQFAAGDLPAPAINVDSVTSKSKEDNXYTRHSLIDGINRGTALIGK
 KVLICGYDQKCAEAMKGGARVSTETDIPINALQAMMEGPDVYVEALIGDADY
 VTATNKDITIMEHKRAMDHALISGIGHFNDIEIDMAGLERSCATRYNRVQNDLMT
 GDTGRSITVLSGRLNLSNATGHSFVSNSNANOTIADIEMTKNDIEDNEYRLP
 KHLDEVARIHVEALGGLTKLTKREQAEYLGVDYEGPKDPDHYK"
 complement(9212..9904)
 /gene="MT3347"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAK47689.1"
 /db_xref="GI:13883167"
 /translation="MPGLKCGKYFRDGGGSLVSTPSATVAPYKRIPIYASRALLR
 DSVLDMRDLTLTRMSAITLSDVARAGISKOTINEFSROGLAGIATRLADRV
 DNVAHSLDANVNEAFVLOGFRSPFAESADPLVLSLGVAPDLQILITTSAPI
 ITRASARLAPADVWATTDNDANVLSRAIVRLCLSYSMPEADHDVADLARLIT
 PFAERHGVINVP"
 complement(9844..10026)
 /gene="MT3348"
 /note="similar to SP:P42453 PID:599606; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="rubredoxin"
 Query Match 36.6%; Score 535.4; DB 1; Length 18249;
 Best Local Similarity 61.3%; Pred. No. 1e-112;
 Matches 879; Conservative 0; Mismatches 551; Indels 3; Gaps 1;
 gene
 CDS
 32 tggagtaacgaattgcgcacatcaactccatgtcttcgagcgaagaactacccttg 91
 9052 TCGACTTTAAGATCGCGACCTGTCACTAGCGGATTCGCGCAAAAGAACTCCGGATCG 8993
 92 ctgaagaagaaatgcagagcttaattgtcttcgtgaagcgttattcgccttaagcaat 151
 8992 CCGAGACGAGATGCCCGGCGCTATGTCGTCGGCGCGAGATATCCGAGGTCGAACCCC 8933
 152 tgaagggtgtcagaatctcgtgtccctccaatgtaacagtcagaagcagcgttcacatg 211
 8932 TGAAGGGGGGCCGATCTCGGGTTCGCTGCACATGACGCGTCCAGACCCCGGGGTGTGATCG 8873
 212 agaacatcaagctctgtgtgtgtatgtcagatgggctctctgcaacatctcttaacg 271
 8872 AAACCTTCACGCGCTGGGGCGCAAGTCCGCTGGGCTCTGTGCAACATCTTCTCCACCC 8813
 272 aagatacagcgtctgtctatcgtlctgcgcccacaacagagacacagaagaagcagccg 331
 8812 AGGATCACGCGCGCGCGCGTGTGATGCGCGCGACAGGCAACCCCGACGAGCCCAAG 8753
 332 gtaaccagctctcgcctgtgaagggcaaacatccacgaatactcgtggagaacacatacc 391
 8752 GTGTCCCGGTGTGCTGTGAAGGGCGAGACGCTGGAAGATTAATGTTGGGGCGCGAC 8693
 392 ggcgtcacatgagcaga--tggtcaaggccacagcaggtgtgtcgtgtgtgtg 448

Db 8692 AGATGCTACAGTGGCCGAGACCCCGCAAGACCCGCCCAACATGATCTCATGACGGCGTG 8633
 QY 449 atgtacactccctcatctccaagggttcgaattcgaaacagccggtgtcttcagagc 508
 Db 8632 ACGCCACCATGTTGGTCTCGCGCGCATGCACTATGACAAAGCCCGGTGTGCGCCG 8573
 QY 509 caacagaagctgaacactcgaatcacgctgtcttctgtcttaactcaagcaggtctca 568
 Db 8572 CCGAGGAGGACGACCCCGCGAGTGGAAAGGCTTCTTCAACCTGCTACGAGACCCGCTTCG 8513
 QY 569 accaagaacaaacactatgcaacacagttgtctgcgagatgaagcgtgttcttcgaagaa 628
 Db 8512 AGACCCACCAAGGACCAATGGACCAAGATAGCCGAGTCCGTCAAGGGCGTCACCGAGGAA 8453
 QY 629 caacaaaggtgtccacgcgcctccacacagcgcgaagaaggaaggaacacacacacacac 688
 Db 8452 CCACACCGCGCTCTCGGTCTACCAATTCGCGCGGCGGGGATCTGGCCTTCGCG 8393
 QY 689 ccatcaacgctcaacagcgtgttaacaaagtcgaagtcgataacactaagcgtcgcgc 748
 Db 8392 CGATCAACGTCACGACTCGGTGACCAAGTCCAAATTCGACAAAGTACGGCCTACCTCGC 8333
 QY 749 actccctatcgtatgtatcaacacggtcttcgagtcgataagtcggcgcaagacagtc 808
 Db 8332 ACTTCCTGATCGAGGATCAACCGCGCGACCGGACCGCGCTGATCGCGGTAAAGAGTCC 8273
 QY 809 tctcaatgggtttacagcgtatgcgcaaggcgtgcgtcaatccctccgttgccaaagcg 868
 Db 8272 TCATCTCGCGCTACCGGAGCTCGGTAAAGGCTGTGCCGAGCGATGAAGGCGCAAGGAG 8213
 QY 869 ctgcggtatcaatcaacgaactcgaaacacacacacacacacacacacacacacacac 928
 Db 8212 CCGGGGCTCTCGTCCGACGAGATCGACCGCATCMACGCGCTGAGGCGCATGAGGAGGCT 8153
 QY 929 accaagtcgc 988
 Db 8152 TCGAGTGGTTCACCTGACGAGAGGCGCATCGGCGACCGGACGACATCGTCAACCGCACCG 8093
 QY 989 gaaactcgatatactctctgttgaatgagtgccagatgaaggaatgaagcgtatgtcg 1048
 Db 8092 GCACCAAGACATCATCTGCTCGAGCATTAAGGCGCATGAAGGACCGCATCTGTG 8033
 QY 1049 gtaacatcgccacttcgataacgaatgttaacagatgagtcctcaatcaatccacagc 1108
 Db 8032 GAATATTCGCGCATCTTCACCAACGAGATCGATGCGCGGCTGGAGCGCTCCGGGCGA 7973
 QY 1109 tcaagcacatcccaatcaagcagaatacagatgtggaattccagatgagccagccta 1168
 Db 7972 CACGGGTCAAGTCAAGCTCAAGTCAAGCTTGACCTTGAGACCTTGAGGCGACGCGCTCGA 7913
 QY 1169 tccctctcttctgtgaagcgcccttcaacatgtgtcgtcgaatgaagtcacacatct 1228
 Db 7912 TCATCTGCTCTTCGAGGCGGCGCTGTGACCTGCGCATCCACCGGCGACCTCTGT 7853
 QY 1229 tccgtatgtcaatgtcaatcaacacagacacacacacacacacacacacacacacacac 1288
 Db 7852 TCGTGTATGACCAAGCTTCGTTACCAACACATGCGCCAGATCGAGCTGTGACCAAGA 7793
 QY 1289 gaggaaatctcgaagaagaagtltaacacatctcgaagatctcgaatgaagaagtcgctc 1348
 Db 7792 ACGACGAGTACGACCAAGAGTGTACCGGCTGCCCAAGACCTCGACGAGAAAGTGGCTC 7733
 QY 1349 gctccacactcgtatctcgaatgtccacacttaacaaagtltaacagaagcagcgtgact 1408
 Db 7732 GAATTCATGTGAGAGCCCTTGGCGGTACCTTACCAAGGTTGACCAAGGAGGAGCGCAAT 7673
 QY 1409 acatcaacggttcaggtgaggttccttaacagtcgtatgtcgtatga 1461
 Db 7672 ACCTCGGCTGAGCTCGAGGCTCCCTTACCAAGCGGACCAACCTACCTACTGA 7620

RESULT 12
 MTY20B11/c

LOCUS MTCY20B11 36330 bp DNA linear BCT 03-AUG-2001
 DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 139/162.
 ACCESSION Z95121 AL123456
 VERSION Z95121.1 GI:3261742
 KEYWORDS
 SOURCE
 ORGANISM Mycobacterium tuberculosis H37Rv.
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;
 Mycobacterium; Mycobacterium tuberculosis complex.
 REFERENCE
 AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eigemeier, K., Gas, S., Barry III, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltham, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moulie, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skellon, S., Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and Barrell, B.G.
 TITLE Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
 JOURNAL Nature, 393 (6685), 537-544 (1998)
 MEDLINE 98295987
 REFERENCE 2 (bases 1 to 36330)
 AUTHORS Parkhill, J.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
 COMMENT On Jun 27, 1998 this sequence version replaced gi:2072692.
 NOTES:
 Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.
 Implemented in Tbpase (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
 CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
 FEATURES
 source
 1. 36330
 location/Qualifiers
 /organism="Mycobacterium tuberculosis H37Rv"
 /strain="H37Rv"
 /db_xref="taxon:83332"
 <1..34116
 /organism="Mycobacterium tuberculosis H37Rv"
 /strain="H37Rv"
 /db_xref="taxon:83332"
 /clone="Y20B11"
 complement(90..848)
 /gene="RV3226c"
 complement(90..848)
 /gene="RV3226c"
 complement(90..848)
 /note="RV3226c" (MTCY20B11.01c), len: 252 aa, unknown, but similar to various hypothetical proteins, e.g. B5D0011_212 Bacillus subtilis O31916 YOWW PROTEIN (224 aa)
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein RV3226c"
 /protein_id="CAB08327.1"
 /db_xref="GI:2072693"
 /db_xref="SPTREMBL:O05872"

misc_feature
 1914..1970
 /gene="aroA"
 /note="PS00885 EPSP synthase signature 2"
 2252..3244
 /gene="RV3228"
 2252..3244
 /gene="RV3228"
 /note="RV3228" (MTCY20B11.03), len: 330. Function: unknown but similar to several hypothetical bacterial proteins, Y160. FASTA best: Y160_HAEIN P45339 Hypothetical protein h11714 (346 aa) opt: 355 z-score: 404.9 E(1): 1.5e-15; (31.7% identity in 281 aa overlap); contains PS00017 ATP/GTP-binding site motif A"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein RV3228"
 /protein_id="CAB08329.1"
 /db_xref="GI:2072695"
 /db_xref="SPTREMBL:O05873"
 /translation="MRPGVDSDVYKVRSSRSPPRTKTPREHADAAMVYVNDGR WGVYGGPRDRTTAMRARELRPIYVGDVYVGLSGPDLTARIVRRPRRTVL RTTADDTPTERTYVAMADLLIVALLADPPRGVLDRLAIIAAGGLTTLCTKT DLAPAPFGKQFDELTVTAAGVDDLLVADILAKITVLLHSGSGKSTLVRLV PADRAVGVETVEITGRGKHTSRVLAIPGLTILSGSGWVITDPCGISFGLAIQPNVL IAFSDIAEATRECPRCGCHMGPPADPECALDITLSPAAARAAARLLAVLSQT"
 2861..2884
 /gene="RV3228"
 /note="PS00017 ATP/GTP-binding site motif A"
 complement(3277..4560)
 /gene="desA3"
 complement(3277..4560)
 /gene="desA3"
 complement(3277..4560)
 /note="RV3229c" (MTCY20B11.04c), len: 427. Function: desA3, but shows some similarity to desaturases. FASTA best: D90914_112 Synecocystis 008671 L1NOLBYOL-COA DESATURASE (EC 1.14.99.25) (359 aa), opt: 319 z-score: 380.8 E(1): 3.2e-14; (25.1% identity in 295 aa overlap)
 /codon_start=1
 /transl_table=11
 /product="desA3"

DEFINITION Sequence 101 from Patent WO0166573.
 ACCESSION AX244109
 VERSION AX244109.1 GI:15859173
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum.
 Corynebacterium glutamicum
 Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
 Actinomycetales: Corynebacteriaceae: Corynebacteriaceae;
 Corynebacterium.
 1 (bases 1 to 1396)
 Pompejus, M., Kroege, B., Schroeder, H., Zelder, O., Habermeyer, G.,
 Kim, J. W., Lee, H. S. and Hwang, B. J.
 Corynebacterium glutamicum genes encoding metabolic pathway
 proteins
 Patent: WO 0166573-A 101 13-SEP-2001;
 BASF AKTIENGESELLSCHAFT (DE)
 FEATURES
 source
 1. 1396
 Location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /db_xref="taxon:1718"
 101..21396
 /note="FRX01371"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC88533.1"
 /db_xref="GI:15859174"
 /translation="MAOVDFKVALDLSLAEGRHRIARAEYEMPGMLQKREFADDEP
 LKARIASIHHTVOTAVILETLALGAVERASCNIESPDDEMAAIIWVSGTVEEP
 AGVDFANKGESLEIEFWICINQIFSWGDELPNITLDDGDTAAVITGREYQGLVP
 PAEANDSEYIAFLGMLREVLAEPKWKIATAVAVTEETTVHRLYHFAEGVL
 PEPANVNDVAIVSKEDNRYGTSHSLVIDINRATDMLGKKNLVCGYGVGKCAEA
 FPGOGARVETVADPINALQALMDGYVTVDEAIEDADIVITATGKMDIISFQOMK
 MKDHALINIGHDEIDHMSLHRDVTRTTKPQVDETFSTGRSIIIVSEGRLLN
 LGMNTHPSFVMSNSFADQTIQIIEIFQMGVEYENETRL"
 BASE COUNT 309 a 396 c 384 g 307 t
 ORIGIN
 Query Match 36.2%; Score 529.6; DB 6; Length 1396;
 Best Local Similarity 64.3%; Pred. No. 2e-111;
 Matches 828; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 512 cagaagctgacaacccctcgaatccgctgcttctgtacactcaagcagcttcaacc 571
 DB 588 AGGCCACAGATTTCCGATGAGTACATCCGATCTTGGGCACTCGCGTAGGTTCTGGCT 647
 QY 572 aagaacagaacacacacacagctgtctcgcgcatgaacagctgttccgaagacaa 631
 DB 648 CAGAGCCTGGCAAGTGGGGCAAGATCGCTGAGCGCGTTAAAGGTGTCAACGAGAAACCA 707
 QY 632 caacaggtgtccacccgcttaccagctcgaagaggggaaactcctctccgaacca 691
 DB 708 CCACCGGTGGCACCCGCTTACCATCTTCGTGAAGAGGGGCTGCTTCCAGCGA 767
 QY 692 tcaagctcaacagcagctgttacaagaagtcgaagtcgaatacctcagctgcgcacat 751
 DB 768 TGAAGCTCAAGAGAGCGCTGTACCAAGTCAAGTGAACAGTAGGACCGACCCGCACT 827
 QY 752 cccctacgagtgtatcaaccgctgtctcgaatgtcaatgtatcggcggaagacgctcgc 811
 DB 828 CCTGTATGACAGCGCATCAACCGCGCCATGCAATGCTCATGGCGGCAAGACGTTTG 887
 QY 812 tcatgggttaacggcagatgtgcgaagagctgcgctcaactccctcgttgccaagcgctc 871
 DB 888 TCTGCGGTTCAGCGGATGTGCGCAAGGGCTCGCTGAGGCTTTTCAGCGCCAGGGCGCTTC 947
 QY 872 gcggtatcatcaagaac 931
 DB 948 GCGTCAAGGTCACCGAAGTCAGACCAATCAACGCTCTTCAAGGCTGTGATGATGCTACT 1007
 QY 932 aggtccgcgcac 991
 DB 1008 CTTGTGTACCGTTGATGAGGCGCATTCGAGGAGCGCGGACATCGTATCACCGGACCGGCA 1067
 QY 992 acgcgatacatcctcgttgacatgtatgcacacacacacacacacacacacacacacac 1051
 DB 1068 ACAAGGACATCATTTCTTTCAGGACATGCTCAAGATGATGATGATGATGATGATGATGAT 1127
 QY 1052 acatcgccac 1111
 DB 1128 ACATCGCTCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187
 QY 1112 agac 1171
 DB 1188 CCCGACACACATCAAGGAC 1247
 QY 1172 tccctctgttgagggcgccctcttcaacacacacacacacacacacacacacacacac 1231
 DB 1248 TCGNCTGTGCGAAGGTGCGCTGTGTAACCTTGCAACCGGACCCACCATCATTTTG 1307
 QY 1232 ttatgtcaatgtatcacaac 1291
 DB 1308 TCAATGTCACCTTTTGCGCATGACACATTTGCGAGATGACATGTTCCAAACGAAAG 1367
 QY 1292 gaatcgcgaagaaggtttacacact 1319
 DB 1368 GACAGTACGACAAAGAGGTCTACCGTCT 1395
 RESULT 15
 WHTSHH 1708 bp mRNA linear PLN 27-APR-1993
 LOCUS Triticum aestivum S-adenosyl-L-homocysteine hydrolase (SH6.2) mRNA,
 DEFINITION complete cds.
 ACCESSION L11872
 VERSION L11872.1 GI:170772
 KEYWORDS S-adenosyl-L-homocysteine hydrolase.
 SOURCE Triticum aestivum (cultivar Warigal) root-tip cDNA to mRNA.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1708)
 REFERENCES
 AUTHORS Richards, K.D. and Gardner, R.C.

Mon Aug 19 15:16:23 2002

us-09-759-990-1.rge

Page 20

Job time: 12776 sec
